U.S. DEPARTMENT OF COMMERCE
National Technical Information Service
PB80-158108

Sediment-Pollutant Relationships in Runoff from Selected Agricultural, Suburban, and Urban Watersheds A Statistical Correlation Study

Tetra Tech, Inc., Lafayette, CA

Prepared for

Environmental Research Lab., Athens, GA

Jan 80

United States
Environmental Protection
Agency

Environmental Research
Laboratory
Athens GA 30605

EPA-600/3-80-022
January 1980

PB80-158108

Research and Development

EPA

Sediment-Pollutant Relationships in Runoff from Selected Agricultural, Suburban, and Urban Watersheds

A Statistical Correlation Study

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ECOLOGICAL RESEARCH series. This series describes research on the effects of pollution on humans, plant and animal species, and materials. Problems are assessed for their long- and short-term influences. Investigations include formation, transport, and pathway studies to determine the fate of pollutants and their effects. This work provides the technical basis for setting standards to minimize undesirable changes in living organisms in the aquatic, terrestrial, and atmospheric environments.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-600/3-80-022

2.

3. RECIPIENT'S ACCESSION NO.
PB80-158108

4. TITLE AND SUBTITLE
Sediment-Pollutant Relationships in Runoff from Selected Agricultural, Suburban, and Urban Watersheds:
A Statistical Correlation Study

5. REPORT DATE
January 1980 issuing date

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Stanley W. Zison

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Tetra Tech, Incorporated
3700 Mt. Diablo Blvd.
Lafayette, California 94549

10. PROGRAM ELEMENT NO
A34B1B

11. CONTRACT/GRANT NO.
68-03-2611

12. SPONSORING AGENCY NAME AND ADDRESS
Environmental Research Laboratory--Athens GA
Office of Research and Development
U.S. Environmental Protection Agency
Athens, Georgia 30605

13. TYPE OF REPORT AND PERIOD COVERED
Final, 9/77-9/78

14. SPONSORING AGENCY CODE
EPA/600/01

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Data from agricultural, suburban, and urban watersheds were subjected to statistical correlation analysis to estimate potency factors. These factors are coefficients that, when multiplied by sediment mass emission rates (transported in runoff), provide estimates of mass emission rates for other pollutants. The potency factors are required input for such lumped-parameter runoff models as the Nonpoint Source (NPS) Model and the Stormwater Management Model (SWMM).

The data were also subjected to multiple regression analysis to examine the effect of storm parameters on runoff water quality and the interrelationship among runoff water quality constituent concentrations themselves (other than sediment load). The multiple regression analysis was primarily exploratory with the objectives of explaining variance of water quality and identifying important independent or predictor variables rather than developing predictive expressions.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Simulation Water Quality Runoff	Model Studies Nonpoint Pollution Agricultural Runoff Agricultural Watersheds Urban Runoff	02A 08H 12A 68D

18. DISTRIBUTION STATEMENT
RELEASE TO PUBLIC

19. SECURITY CLASS (This Report)
UNCLASSIFIED

20. SECURITY CLASS (This page)
UNCLASSIFIED

21. NO. OF PAGES

22. PRICE

EPA Form 2220-1 (9-73)

☆ U.S. GOVERNMENT PRINTING OFFICE 1980-657-146/5566

EPA-600/3-80-022
January 1980

SEDIMENT-POLLUTANT RELATIONSHIPS IN RUNOFF FROM SELECTED AGRICULTURAL, SUBURBAN, AND URBAN WATERSHEDS

A Statistical Correlation Study

by

Stanley W. Zison
Tetra Tech, Inc.
Lafayette, California 94549

Contract No. 68-03-2611

Project Officer

Charles N. Smith
Technology Development and Applications Branch
Environmental Research Laboratory
Athens, Georgia 30605

ENVIRONMENTAL RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U. S. ENVIRONMENTAL PROTECTION AGENCY
ATHENS, GEORGIA 30605

DISCLAIMER

This report has been reviewed by the Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, Georgia, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

FOREWORD

As environmental controls become more costly to implement and the penalties of judgment errors become more severe, environmental quality management requires more efficient analytical tools based on greater knowledge of the environmental phenomena to be managed. As part of this Laboratory's research on the occurrence, movement, transformation, impact, and control of environmental contaminants, the Technology Development and Applications Branch develops management or engineering tools to help pollution control officials achieve water quality goals through watershed management.

Essentially all control technology for nonpoint sources is related to erosion or sediment control. Several techniques are available today that estimate nonpoint source pollutant loadings based on the assumption that a reasonable correlation exists between various pollutants and sediment. For these tools to be useful, sediment-pollutant correlations (potency factors) must be developed to provide essential inputs into lumped-parameter runoff models such as the EPA's Nonpoint Source Model (NPS) and Storm Water Management Model (SWMM). This report provides statistical correlation values for the potency factor, based on estimates from runoff data representing agricultural, suburban, and urban watersheds.

David W. Duttweiler
Director
Environmental Research Laboratory
Athens, Georgia

ABSTRACT

Data from agricultural, suburban, and urban watersheds were subjected to statistical correlation analysis to estimate potency factors. These factors are coefficients that, when multiplied by sediment mass emission rates (transported in runoff), provide estimates of mass emission rates for other pollutants. The potency factors are required input for such lumped-parameter runoff models as the Nonpoint Source (NPS) Model and the Stormwater Management Model (SWMM).

The data were also subjected to multiple regression analysis to examine the effect of storm parameters on runoff water quality and the interrelationship among runoff water quality constituent concentrations themselves (other than sediment load). The multiple regression analysis was primarily exploratory with the objectives of explaining variance of water quality and identifying important independent or predictor variables rather than developing predictive expressions.

This report was submitted in fulfillment of Contract No. 68-03-2611 by Tetra Tech, Incorporated, under the sponsorship of the U.S. Environmental Protection Agency. The report covers the period September 20, 1977, to September 19, 1978, and work was completed as of September 19, 1978.

CONTENTS

FOREWORD iii
ABSTRACT iv
FIGURES vi
TABLES viii
ACKNOWLEDGEMENTS xi

1. INTRODUCTION 1
2. CONCLUSIONS AND RECOMMENDATIONS. 3
    CONCLUSIONS 3
    RECOMMENDATIONS 4
3. BACKGROUND AND TECHNICAL ISSUES. 6
    SUSPENDED SOLIDS AND WATER QUALITY. 7
    PAST STUDIES. 10
4. METHODOLOGY. 28
    GENERAL ANALYSIS PROCEDURE. 28
5. RESULTS AND DISCUSSION 51
    SIMPLE LINEAR REGRESSION. 51
    MULTIPLE REGRESSION 75

REFERENCES 112
APPENDIXES

A. REGRESSION ANALYSIS THEORY 114
B. RUNOFF WATER QUALITY ANALYSIS METHODS. 126

FIGURES

Number Page

1 Nitrogen availability pattern in fresh urban runoff sample(s) 11

2 Pollutant variations with Q and time for storm no. 13, date 3/16/72 14

3 Pollutant variations with Q and time for storm no. 13, date 3/16/72 15

4 Pollutant variations with Q and time for storm no. 20, date 6/20/72 16

5 Pollutant variations with Q and time for storm no. 20, date 6/20/72 17

6 General data analysis procedure 29

7 Sample plot showing output of simple linear regression program used in this study 31

8 Soils and topography, watershed P-01 40

9 Soils and terrace configurations, watershed P-03 41

10 Soils and terrace configurations, watershed P-04 42

11 Buffalo Bill Watershed configuration and location 44

12 Location and configuration of Michigan State University test plots 47

13 Location map for Seattle catchments 49

14 A comparison of statistical results among Southcenter, South Seattle, and Viewridge 1 70

15 Scatter plot and regression of iron concentration on suspended solids concentration 73

16 Scatter plot of $\log_{10}$ iron concentration on $\log_{10}$ suspended solids concentrations 74

FIGURES (continued)

Number Page

A-1 Schematic showing an interpretation of R^2 118

TABLES

Number		Page
1	Pollutant Fractions Associated with Particle Sizes	13
2	Total Suspended Solids and Organics Data Averaged for 36 Storm Events and Ratios of Organics to Total Suspended Solids	18
3	Total Phosphorus, Kjeldahl Nitrogen, and Fecal Coliforms in Runoff Averaged for 36 Storm Events and Ratios of Each to Total Suspended Solids	20
4	Metals Concentrations (Ca, Co, Cu, Cr, Fe) Averaged for 36 Storm Events and Ratios of Each to Total Suspended Solids	22
5	Metals Concentrations (Pb, Ni, Mg, Mn, Zn) Averaged for 36 Storm Events and Ratios of Each to Total Suspended Solids	24
6	Regression of Log [NO_3 + NO_2]	33
7	Significance of Entries in Table 6	35
8	Summaries for Data Bases Used in this Study	39
9	Buffalo Bill Watershed Flow and Rainfall Data	45
10	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From Watkinsville, Georgia - Data are From Plot P-04	52
11	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From Watkinsville, Georgia Test Plot P-03.	53
12	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown. Data are From Watkinsville, Georgia Test Plot P-01	54
13	Regression of Dissolved Species on Concentration in Suspended Solids Using Data From Watkinsville, Georgia	56
14	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From the Buffalo Bill Watershed, Iowa	58

TABLES (continued)

Number		Page
15	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From Michigan State University Study	60
16	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From Viewridge 1, Seattle Data	61
17	Statistics for Correlations Between Turbidity and Constituents Shown From Viewridge 1, Seattle Data	63
18	Statistics for Correlations Between Antecedent Dry Days and Constituents Shown From Viewridge 1, Seattle Data	64
19	Statistics for Correlations Between Suspended Solids Concentrations and Constituents Shown From the South Seattle Data Base	66
20	Statistics for Correlations Between Antecedent Dry Days and Constituents Shown From South Seattle Data Base	68
21	Statistics for Correlations Between Suspended Solids and Constituents Shown From Seattle Southcenter Data Base	69
22	Statistics for Correlations Between Suspended Solids and Constituents Shown From the Honey Creek Watershed	72
23	Summary Listing of Potency Factors Estimated From Agricultural, Suburban, and Urban Watersheds Examined in This Study[A]	76
24	Variables Examined in Multiple Regression as Candidate Predictors ofRunoff Water Quality	78
25	Simple Correlations (r) Between Dependent and Independent Variables for Watkinsville Plot P-04 Data	79
26	Multiple Regression Statistics forWatkinsville Plot P-04	81
27	Regression Equations With Confidence Intervals for Slope and Intercept From Watkinsville Plot P-04	82
28	Simple Correlations (r) Between Dependent Variables (Trifluralin and Diphenamid) and Independent Variables Data are From Watkinsville Test Plots	84
29	Multiple Regression Statistics for Watkinsville Plots P-01 and P-04 Data	86
30	Regression Equations with Confidence Intervals for Slope and Intercept From Watkinsville Plots P-01 and P-03 Herbicide Data . .	87

TABLES (continued)

Number | Page

31 Simple Correlations (r) Between Dependent and Independent Variables for the Buffalo Bill Watershed 89

32 Multiple Regression Statistics for the Buffalo Bill Watershed . . . 90

33 Regression Equations with Confidence Intervals for Slope and Intercept From the Buffalo Bill Watershed Data Base 91

34 Simple Correlations (r) Between Dependent and Independent Variables for Southcenter, Seattle 94

35 Multiple Regression Statistics for Southcenter, Seattle 95

36 Results of Validation-Generalization Using Data From Southcenter and Coefficients From Viewridge 1 97

37 Regression Equations with Confidence Intervals for Slope and Intercept From Southcenter, Seattle Data 98

38 Simple Correlations (r) Between Dependent and Independent Variables (South Seattle) 100

39 Multiple Regression (t) Statistics and R^2 for South Seattle Data 102

40 Results of Validation-Generalization Using Data From South Seattle and Coefficients From Viewridge 1 103

41 Regression Equations with Confidence Intervals for Slope and Intercept From South Seattle Data 104

42 Simple Correlations (r) Between Dependent and Independent Variables (Viewridge 1) 107

43 Multiple Regression (t) Statistics and R^2 for Viewridge 1 (Seattle) Data 109

44 Regression Equations, with Confidence Intervals for Slope and Intercept From Viewridge 1 (Seattle) Data 110

ACKNOWLEDGEMENTS

The author would like to express his gratitude to the Environmental Research Laboratory at Athens, Georgia, for its support of this study (EPA Contract Number 68-03-2611) and in particular, to Messrs. Charles N. Smith and Rudolf Parrish for their suggestions and editorial comments. Thanks also go to Messrs. William B. Mills, Paul Johanson, Larry Woods, and Mrs. Mei-Chi Hua for assisting with analyses; to Mrs. Bernice Bujacich and Ms. Susie Madson for typing; and to Mrs. Anna Zison for drafting.

Finally, the author would like to thank Dr. Dave Baker (Heidelberg College, Tiffin, Ohio), Mr. David M. Cline (EPA-ERL, Athens, Georgia), Mr. Roger Splinter (Iowa State Hygienic Laboratory, Iowa City), and Dr. Wayne Huber (University of Florida, Gainesville) for providing data.

SECTION 1

INTRODUCTION

In recent years, much attention has been paid the problem of point discharges, their effects on receiving waters, and the study and simulation of resulting water quality interactions. Less work has been done on the problem of nonpoint loadings, despite the fact that stormwater runoff is frequently the most significant determinant of receiving water quality. Stormwater runoff can be important, for example, in rural areas where point sources are minor and where intensive farming and other agricultural activities (e.g., feed-lot operation) lead to heavy concentrations of potential pollutants spread over the land. In large urban areas, where point sources are commonly very important, nonpoint loads may also be substantial as they contribute litter, heavy metals, gardening wastes, and construction debris through the storm sewers to the receiving waters.

Emphasis has commonly been placed on the impacts of point sources rather than on nonpoint sources because:

- Point discharges tend to be easier to identify than nonpoint, and therefore draw public attention. Further, this has brought them to the fore among environmental researchers.
- Point discharge characteristics tend to be easier to quantify and are commonly less variable over time than are nonpoint characteristics.
- Reliable mechanistic predictive tools have been easier to develop where only point sources are considered, since in the overall modeling process, the point source represents a boundary condition, while nonpoint loads must themselves be estimated. This means one more level of simulation--and one of substantial difficulty and uncertainty.

The combined impact of the importance of nonpoint wasteloads and the difficulties inherent in modeling their effects has led to development of various storm runoff models such as the Environmental Protection Agency's (EPA) Stormwater Management Model (SWMM) (1), EPA's Nonpoint Source Pollutant Loading Model (NPS) (2), STORM (Water Resources Engineers/U.S. Army Corps of Engineers) (3), EPA's Agricultural Runoff Management (ARM) Model (4), and EPA's HSP-F (5).

With respect to runoff water quality, these models, and others of this class, may be categorized into those which attempt to mechanistically simulate soil and runoff water chemistry, and those which assume the mass emission rate of each pollutant to be a simple function of sediment mass emission rate in runoff from the watershed. ARM falls into the first category (4). It includes algorithms for simulating pesticide decay on the watershed such that the availability of these pollutants to be washed off during a storm is a function of time elapsed since application. Analogously, for nutrients, the ARM model simulates soil storage and decay between storms, thereby mechanistically estimating quantities of pollutants available for washing off in each storm being modeled.

In contrast, SWMM, STORM and NPS all simulate runoff water quality simply as a linear function of total sediment transported, with appropriate multipliers (potency factors) provided as input. The representations are of the form:

$$C = P \cdot S \qquad (1)$$

where C = mass of pollutant washing off the watershed, any mass units

P = potency factor converting S to C, units depending on units of S and C. For use with the NPS model, P is expressed as a percent. Therefore P, as presented later in this report, must be multiplied by 100.

S = mass of sediment washing off the watershed, any mass units

Recognizing that SWMM, NPS, and other similar models are popular and are likely to continue in common use, EPA's Environmental Research Laboratory at Athens, Georgia, funded a statistical study geared essentially toward defining the nature of relationships between runoff sediment loads and concentrations of other water quality constituents. This report discusses results of that statistical study and emphasizes observed correlations between runoff water quality and suspended sediment. Secondarily, findings are also presented which relate runoff water quality to storm characteristics and certain other variables. The former provides estimated potency factors for use with such models as SWMM and NPS, and the latter can provide guidance in further model and research study development.

SECTION 2

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

1. The temporal variance of suspended sediment concentration in storm runoff can account for a relatively small proportion of the temporal variance of nearly all other water quality constituents considered.

2. Watershed type appears to be very important in determining the reliability of potency factors. Based upon the results of this study, potency factors computed for urban runoff are more reliable than those developed for suburban, rural, and agricultural areas.

3. There is a very substantial variability of potency factors among agricultural watersheds representing diverse geographical and climatological regions.

4. Within the single urban area examined (Seattle) the potency factors are quite similar from sampling station to station for a number of pollutants. This suggests possible similarities among pollutant sources, and deposition and transport phenomena as well.

5. When factors in addition to suspended sediment load are taken into consideration, explication of runoff water quality improves significantly. Factors which proved important include number of dry days preceding storms, time elapsed since the beginning of storms, cumulative rainfall since the beginning of storms (the latter two sometimes simultaneously important), time since application of chemical (agricultural watersheds), and elapsed time since some reference date (suggesting longterm trends).

6. The fact that time since the beginning of storms is sometimes significant even after cumulative storm rainfall is accounted for suggests that wetting rates may be important in the temporal profile of runoff water quality from a watershed. Additionally, it may be that some pollutants do not reside at the very surface, and some amount of surface material must be washed off before those pollutants are exposed to the dislodging and transporting action of raindrop impingement and overland flow.

7. In some cases, the correlation between suspended solids and dry days preceding storms is relatively weak. In many of these same cases, the correlation between suspended solids and certain water quality constituents is also weak, while the direct correlation between dry days preceding storms and runoff water quality concentrations is substantially stronger.

Both SWMM and the NPS model simulate dust and dirt accumulation as a function of dry days preceding storms, and compute water quality as a function of sediment transported.

8. Correlations among water quality constituents are generally strong relative to other correlations examined. In some cases, however, correlations are lower than might be expected (e.g., suspended solids concentration with turbidity).

9. Rainfall intensity, lagged 0, 1, and 2 hours prior to the time of water quality sampling, is a consistently poor predictor of runoff water quality. This may be due simply to inadequate data, however, rather than to a lack of real relationships between rainfall intensity and runoff water quality. Data to examine temporally lagged rainfall intensity were severely limited in this study.

10. In the development of the potency factors reported here, a simple linear model was fitted to the available data. This model included a constant (y-intercept) which was not forced to zero. Also, in many instances fitting the linear model without transforming the data would have resulted in clustering of points, unacceptable weighting of selected data, and biasing of computed statistics. Transforming the data to a log scale was generally found to be an acceptable solution. However, both the existence of a non-zero y-intercept and the use of a log transformation mean that some of the potency factors reported here are only approximations to those required by such models as SWMM and NPS.

RECOMMENDATIONS

Modeling nonpoint pollution transported in storm runoff is a difficult task at best, based upon the findings of this and innumerable other studies. The difficulties stem, in part, from the complexities of sediment and pollutant transport phenomena. Perhaps even more important, however, is the difficulty of satisfactorily describing, in mathematical or modeling terms, any real watershed of significant size. It would be easy to recommend, based upon the results obtained in this study, that much further research be done to characterize pollutant deposition and transport phenomena, especially in terms of accounting for the spatially variable nature of potency factors. Yet, it seems clear that the costs of an effective program would be very large, particularly for such studies of non-urban watersheds. In such watersheds, there are almost certainly too many interdependent factors involved which influence runoff chemistry. These have too broad a range of values from watershed to watershed to make practical a few comprehensive pollutant washoff studies of universal validity. Considering the large and small scale complexities of real watersheds, the interaction of factors affecting runoff water quality, and the fact that the impact of watershed topographic features is strongly dependent upon the spatial orientation and positioning of those features, the model resolution needed to realistically represent a watershed spatially seems prohibitive.

It appears, then, that modelers must be content to represent large watersheds in the grossest of terms using lumped-parameter models, and to consider only spatially and temporally averaged phenomena. Given that such is the case, funding agencies and the research community must decide just where advances in the state of the art can realistically be expected. The primary recommendation of this study is that the real need for accuracy in nonpoint waste load estimation be carefully assessed and that future directions in nonpoint modeling be very carefully planned. As far as potency factors are concerned, this would seem to emphasize urban watersheds, followed by suburban, rural, and agricultural watersheds. Further, the practicality, value, and cost of developing improved models should be weighed against their potential benefits, the costs of field sampling and direct watershed characterization on an individual basis, and the acceptability of statistical, empirical, or simplified computational procedures for nonpoint loading estimation.

Specific recommendations regarding potency factors as input to the NPS model or SWMM are as follows:

1. In applying potency factors to non-urban watersheds, modelers should be especially cognizant of the site-specific nature of the potency factors as well as the large degree of uncertainty in the values. The latter is indicated by various statistics presented in the text.

2. Where the y-intercept is non-zero, either the particular model should be altered to accommodate this or at least the possibility of error introduced by ignoring the constant should be recognized. Where the potency factor has been developed using log-transformed data, the model might also be modified, or a piecewise-linear approximation might be used.

3. In view of some of the findings outlined above under "Conclusions," item 7, the modeler might consider altering the code for selected constituents. In some cases (see pertinent parts of the text) it may be appropriate to model water quality directly as some function of antecedent dry days. Although this is less appealing conceptually than predicating pollutant concentrations on sediment loads, correlation results suggest it may allow for greater accuracy.

4. Modelers should recognize that the fundamental assumption underlying the concept of potency factors is some reasonably consistent relationship between quantity of sediment transported and concentration of other constituents. Even if one ignores differences in particle sizes, suspended sediment surface area-to-mass ratios, sediment composition, and the fact that a given pollutant may not, in fact, be physically associated with particulate, and even if one assumes such a relatively consistent relationship between suspended sediment and pollutant concentration, it is still possible that there may be little or none of the pollutant available to wash off. At the other extreme, the pollutant may be present in such amounts as to swamp available sediment sorption sites. The result is a large error in the potency factor even under such ideal assumptions. The modeler, therefore, should recognize the conceptually tenuous nature of the potency factor itself, and use it accordingly.

SECTION 3

BACKGROUND AND TECHNICAL ISSUES

The chemical characteristics of stormwater runoff are influenced by many different phenomena which are manifested from raindrop nucleation in the atmosphere to the point where the droplet ultimately intercepts a receiving water body. The phenomena include:

Scrubbing of atmospheric pollutants,

Surface impingement and suspension of particulate matter,

Dissolution of chemical species during and following impingement on a surface, and

Chemical transformations in the aqueous phase.

Atmospheric scrubbing determines the chemical characteristics of the raindrop to the point of impaction with the earth's surface. Pollutants washed out and contained within the raindrop can include particulate matter, sulfuric acid, nitrate and nitric acid, complex organic substances, traces of hydrocarbons, and other substances introduced by local atmospheric conditions. Upon and following surface impingement, a vast array of substances is available for transport. Much of it is mineral (clays, silts), while a substantial portion is organic. Depending upon the site of impaction and the path of runoff, material transported may include significant amounts of animal wastes, litter, chemicals of anthropogenic origin, natural atmospheric fallout, soil particles, and plant debris. In transit to a receiving water channel, chemical and physical transformations also occur. Although biochemical decay is generally considered unimportant given the time scale involved, physical effects, such as disaggregation of clumps of matter, sorption, and chemical phenomena (e.g., complexation, precipitation, and dissolution) can all be important determinants of runoff water quality.

This study, being a statistical rather than mechanistic investigation, and using actual field data, directly or indirectly involves virtually all of the phenomena which affect runoff water quality. The study specifically examines relationships between sediment loads and runoff water quality, which is the major emphasis, and then examines certain other variables (e.g., time since beginning of storm, cumulative rainfall) hypothesized to influence the chemical composition of the runoff.

SUSPENDED SOLIDS AND WATER QUALITY

Potency Factors

Potency factors are used in the NPS and SWMM models as described in equation 1. In later sections of this report, P is denoted as m, which is a computed approximation to the true value of P. It is based upon concentrations of pollutant and sediment in storm runoff rather than actual mass emission rates. The way in which "S" in equation 1 is defined is important in interpreting the meaning of potency factors. Commonly, in applications of the models to real watersheds, mass emission data are unavailable, and the modeler finds himself calibrating with data representing total concentrations (best), concentrations of dissolved species, or possibly even concentrations in the suspended solids fraction. The significance of any potency factor depends upon which of these is being used. If, for example, the calibration is made using dissolved lead values, there are two major determinants of the potency factor values. One is the partition coefficient for lead between suspended solids and the aqueous phase. The other is the amount of lead available to wash off the watershed relative to the amount of runoff transporting the lead. Thus, where the partition coefficient favors the solid phase and there is little lead to be washed off, the value of P in equation 1 for dissolved lead could be extremely low. It is clear that an increase in the amount of available lead, a change in the partition coefficient favoring lead in the aqueous phase, or both, would increase P. The same is true if the suspended fraction lead concentration were of interest, only with the change in the partition coefficient favoring the suspended sediment. In contrast, however, if total lead is used to estimate P, the partition coefficient is irrelevant, while the amount of lead available on the watershed to wash off still influences P.

The fact that the amount of any substance available to wash off strongly affects P means that P must inherently be only as consistent as is the rate of pollutant accumulation on the watershed. Since it is clear that the rate of accumulation of many pollutants on watersheds is quite variable (e.g., nutrients and pesticides in agricultural lands, salt on city streets) P for those pollutants must similarly be quite variable.

With respect to the *reliability* of simulations of suspended or aqueous fraction pollutant levels, the partition coefficient is important to the extent it is variable. As will be described later, studies have suggested that the partition coefficient is quite variable for many kinds of pollutants including nutrients.

Given these sources of variability, and where P is to be estimated from field data, P for dissolved species might be considered conceptually as:

$$P_B = \frac{C_B}{p \cdot S \cdot A} \quad \text{or} \quad C_B = P_B \cdot S \cdot p \cdot A \qquad (2)$$

where P_B = potency factor for pollutant B assumed to be 50% in aqueous phase and maximally available to wash off

C_B = concentration of dissolved species B

S = concentration of suspended matter

p = partition coefficient ($C_{sediment}/C_B$) for B distributed between the suspended matter and the aqueous phase.

A = Availability factor, maximum value = 1 where the greatest amount of B has accumulated to wash off. Where none has accumulated, A approaches zero.

The fact that the potency factor is not ordinarily estimated as in equation 2--that is, p and A are unavailable and represent random variables--means that P determined from field data must be proportionately as variable as phenomena represented by A and p. Since A is important regardless of whether dissolved, suspended, or total pollutant is estimated, and since A is likely to be quite variable, P_B computed from the pollutant concentration and the suspended sediment load must be subject to very substantial error as well.

Pollutant Transport

Relationships between suspended solids and water quality involve the following.

- Transport as particulate
- Dissolution, precipitation, and volatilization
- Sorption phenomena

Transport as Particulate--

To varying degrees, water quality constituents of interest may exist as colloidal or larger compositionally homogeneous masses and are then transported as a fraction of the suspended solids. Examples include coliforms, BOD (biochemical oxygen demand), and naturally occurring mineral matter. Additionally, where surface dumping of waste chemicals occurs, virtually any low solubility species may be transported long distances this way.

Dissolution, Precipitation, and Volatilization--

At any point in time, the aqueous portion of runoff tends to approach a state of equilibrium among the rates of dissolution of each chemical species, of removal to the solid phase, and of exchange with the atmosphere. Because conditions are continually changing with location, equilibrium is not likely to be attained for any significant period of time, and the process is one of continuous changes in the rate and direction of approach to concentration equilibrium.

The rate of approach to equilibrium is ordinarily influenced by such factors as the instantaneous concentration of the dissolved form, the availability and nature of the substance in the solid phase and the particle sizes involved, the volatility of the substance, the availability of other chemical species and ions with which a precipitate may form, the turbulence and rate of flow and resulting rates of mass advection-diffusion, and temperature. The equilibrium concentration itself is mainly dependent upon the particular chemical species involved, the presence of other species, and the pertinent solubility products or stability coefficients. These, in turn, are a function of temperature, but also may be influenced by pH and redox potential.

Precipitation, it should be recognized, is not a simple process of solid matter coming out of solution and settling or being transported in pure form. In natural systems, several phenomena may be of importance with respect to precipitation and transport in the solid phase. These include adsorption (discussed below) and occlusion. Occlusion is the entrapment of water within a precipitate or floc which may result in any substance transported being associated with the suspended matter fraction.

Sorption Phenomena--

The term "sorption" refers to the physical association of a chemical species with particulate matter. Where particulates are of small size and concentrations are high (as in water turbid with clay) such substances as pesticides and heavy metals may be associated with suspended solids because of surface adsorption. Adsorption, it should be noted, may be due to electronic charge, as in ions adsorbed on the surface of a crystal lattice, or due to other kinds of affinities, as in lipophilic substances adsorbing onto organic flocs. As is true for dissolution-precipitation, adsorption in runoff represents a dynamic equilibrium.

Implications for This Study

This study is mainly concerned with the degree to which levels of chemical substances transported in runoff may be correlated with suspended solids loads, and the development of potency factors for use in mathematical modeling. Potency factors, as used in such models as NPS, SWMM, and STORM, may be interpreted as coefficients which when multiplied by the suspended solids concentration, yield concentrations of other species, such as coliforms, nitrate, and BOD. Many studies have provided data with which potency factors can be computed. The potential reliability of each factor, and as a consequence, the reliability of runoff water quality simulations, however, does not depend upon the average degree to which each constituent is associated with suspended matter, but upon the amount of variation in the relationship. Coliforms, for example, might be expected to associate with the suspended fraction. Coliforms themselves are particulate, of course, and if appropriate techniques are used, can always be removed from the aqueous phase. Nevertheless, correlations between coliforms and suspended matter are far from perfect as will be discussed in more detail below. This may reflect errors of measurement, differences in assay technique, nonuniformity of coliform availability for washoff, differences in die-off rates, and a host of other factors.

At the other end of the spectrum, nitrate can be expected to associate with the aqueous phase rather than with the suspended matter, based upon considerations such as the Peneth-Fajans-Hahn Law and the high solubility of virtually every inorganic nitrate species. The Law states that where at least two types of ions are present, and other factors being equal, the ion forming a compound of lowest solubility with the crystal lattice ions will be preferentially adsorbed as will ions with highest charge and those closest in ionic radius to the ion which would normally be at the particular lattice location. With respect to nitrate, correlations may be expected to be poor and erratic, and potency factors are likely to be very unreliable.

Finally, somewhere between the two extremes is BOD. A variable degree of association of BOD with suspended matter can be expected based upon the fact that much BOD can be present as organic particulate, whereas a substantial component may also exist as dissolved organic matter (e.g., organic acids).

A number of studies have been performed in the past, in part geared to determining degrees of association of various water quality constituents (especially nutrients) with suspended matter, although these studies commonly do not have as a specific objective the development of potency factors. In a report on an application of the NPS Model to nutrients in runoff, Donigian and Crawford (6) state that:

> "In summary, the literature appears to indicate sediment can be used as a reasonable indicator of nutrient loadings by surface runoff from agricultural and urban lands. However, soluble nutrient losses can be significant in watersheds where subsurface discharge is a major component of the runoff."

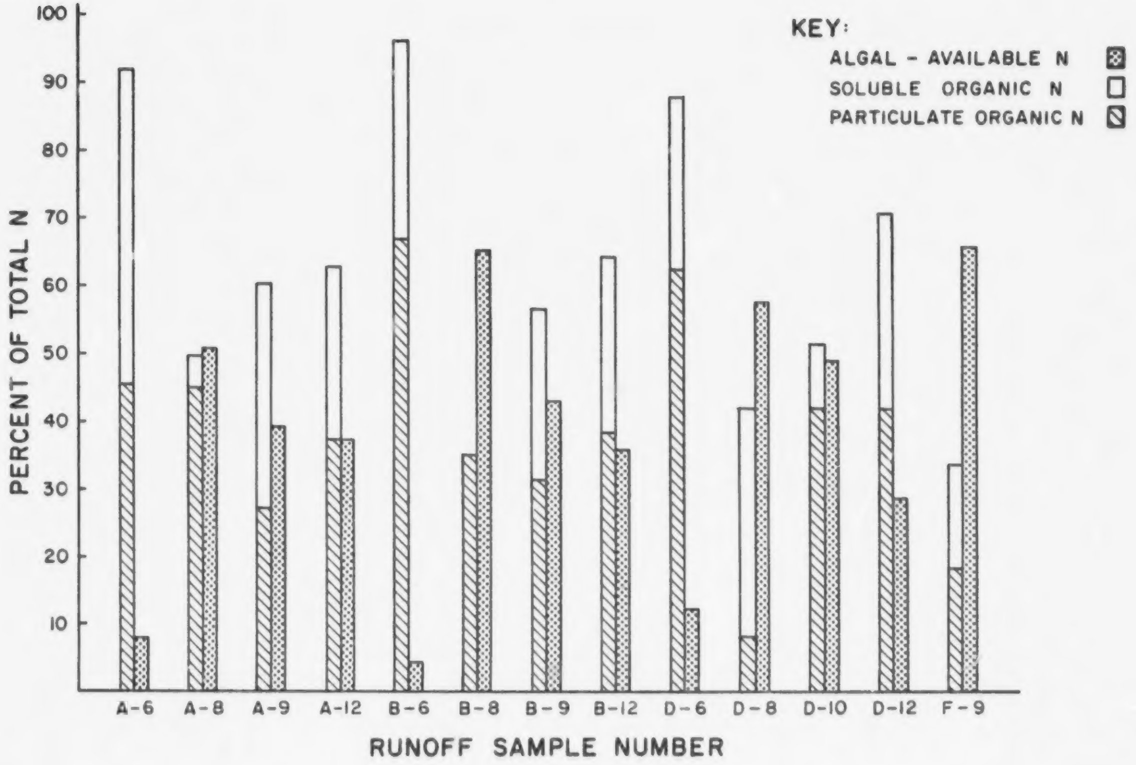
That the literature actually supports the contention that sediment loads are a reasonable indicator of nutrient loads can certainly be debated. Further, the literature suggests that sediment levels are, at best, only a fair index of loads of pesticides, heavy metals, BOD, and other substances which might be expected to associate with particulate matter.

PAST STUDIES

Urban Storm Runoff

Cowen *et al*. (7) investigated nitrogen availability in urban runoff. Findings showed that for organic nitrogen the relative portions present as soluble and as particulate nitrogen varied substantially. Based upon the data represented in Figure 1, from about 0 percent (B-8) to about 80 percent (D-8) of total organic nitrogen is in the aqueous phase. Whereas the samples within a particular letter designation represent different samples from a single site (e.g., A-6, A-8, A-9, A-12 are from site A, a low-density residential area), it is also clear that variability was high within sites (see e.g., A-6 versus A-8 and D-8 versus D-10). Cowen *et al*. (7) point out in general terms, however, that of 13 samples, 10 had more particulate organic nitrogen than soluble organic nitrogen.

It is important to note here that even if the relative proportions of soluble and particulate organic nitrogen were essentially constant, this does not provide information about the ratio of nitrogen to total sediment, which



Site A = low density residential area, B = medium density area, D = University of Wisconsin-Madison campus, F = recently sodded residential area previously under construction.

Figure 1. Nitrogen availability pattern in fresh urban runoff sample(s). (Redrawn after Cowen et al. (7))

is the basis for potency factors. Cowen et al. (7) had not intended to address this issue. The results do suggest, however, that the nature of nitrogen species transported from urban sheds and/or the nature of the sediment tend to be quite variable.

The study by Cowen et al. (7) suggests how variable the partition coefficient can be, and thus has major implications for predicting dissolved or suspended species concentrations. Cowen and Lee (8) have performed a parallel study to that of Cowen et al. (7) examining the availability of phosphorus in urban runoff-transported particulate. Results showed that in the 44 samples taken, particulate phosphorus represented from 13 to 97 percent of total phosphorus.

Sartor et al. (9) have studied runoff quality in 12 U.S. cities having populations ranging from about 13,000 to about 900,000. Table 1 shows data from the study associating various pollutants with particle size ranges. Although Sartor et al. (9) point out the importance of the fine silt-like particulate fraction in that it represents a very small portion of the total solids (TS) but accounts for large portions of other contaminants, they do not discuss the variability of their data within or between the study sites.

Colston (10), in a report on urban land runoff characterization and treatment, has presented data which more directly address the issue of potency factors. Colston monitored storm activity and runoff for 36 storms on the Third Fork Creek Watershed within the City of Durham, North Carolina. Figures 2 through 5 show some of his results. Note that within a given storm event, the curve of total suspended solids (TSS), which is shown on all plots, correlates, at least upon visual inspection, fairly well with certain other variables (notably fecal coliforms, total phosphorus, iron, and lead), but correlates very poorly with some others. For example, the substantial undulations in the curves of TSS do not very well match those of TOC (total organic carbon), BOD (biochemical oxygen demand), or calcium.

In his report, Colston (10) also presents runoff quality for 36 storms averaged by storm. These data, which represent flow at the downstream terminus of the watershed, are presented in Tables 2 through 5. Also shown in the tables are ratios of each pollutant to total suspended solids. The data are representative of the total runoff from the watershed studied since gauging and sampling were done at the downstream watershed terminus. Tables 2 through 5 show the variability of the ratios of each water quality characteristic to suspended sediment--that is, the potency factors. Note in Table 2, for example, that the ratios of 36 storm means of COD, TOC, and BOD to TSS vary by a factor of about 20. Note, also, that concentrations of each of these are substantial, suggesting such variability is not entirely the result of errors of measurement. Similarly, Tables 3, 4, and 5 exhibit high variabilities for ratios of means of the various water quality parameters to the mean of total suspended solids.

With regard to this variability, three additional points should be considered. First, 521 separate samples were taken over the 36 storms, averaging between 14 and 15 samples per storm. Because the ratios shown in Tables 2 through 5 represent ratios of means, variability among the ratios will probably

TABLE 1. POLLUTANT FRACTIONS ASSOCIATED WITH PARTICLE SIZES (9)

Measured Pollutant	Fraction of Total (% by weight)		
	<43 μm	43 μm-246 μm	>246 μm
TS	5.9	37.5	56.5
BOD_5	24.3	32.5	43.2
COD	22.7	57.4	19.9
VS	25.6	34.0	40.4
Phosphates	56.2	36.0	7.8
Nitrates	31.9	45.1	23.0
Kjeldahl nitrogen	18.7	39.8	41.5
All heavy metals	51.2		48.7
All pesticides	73		27
PCB	34		66

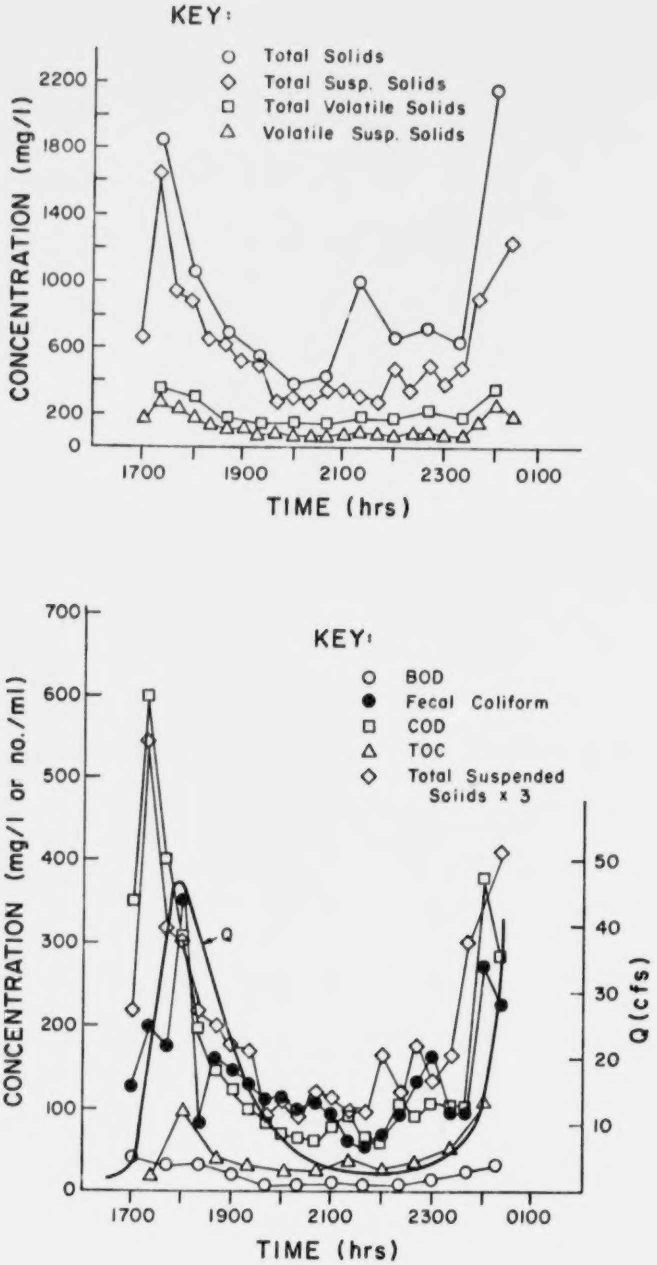


Figure 2. Pollutant variations with Q and time for storm no. 13, date 3/16/72. (Redrawn after Colston (10))

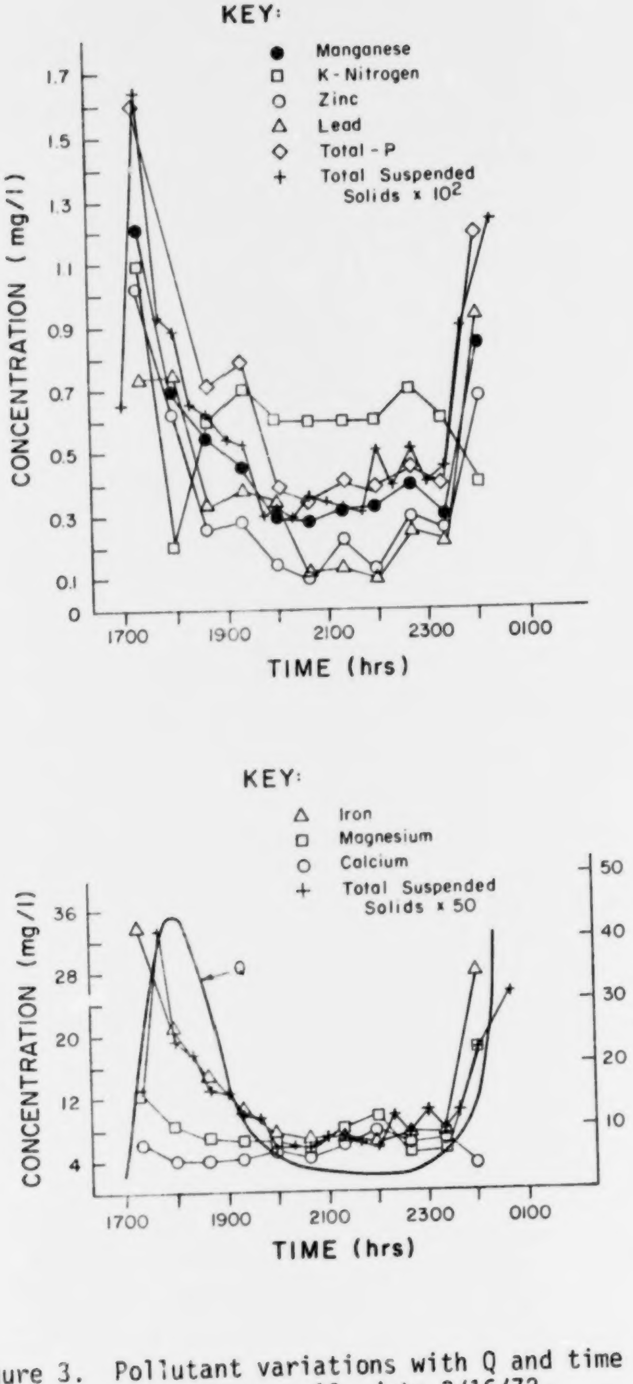


Figure 3. Pollutant variations with Q and time for storm no. 13, date 3/16/72. (Redrawn after Colston (10))

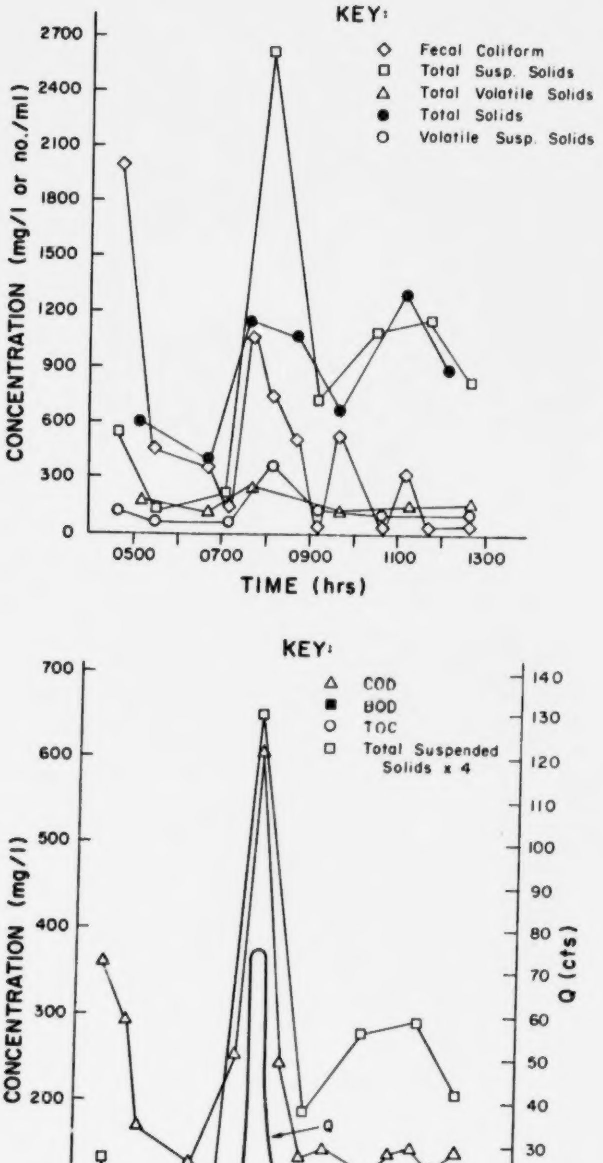


Figure 4. Pollutant variations with Q and time for storm no. 20, date 6/20/72. (Redrawn after Colston (10))

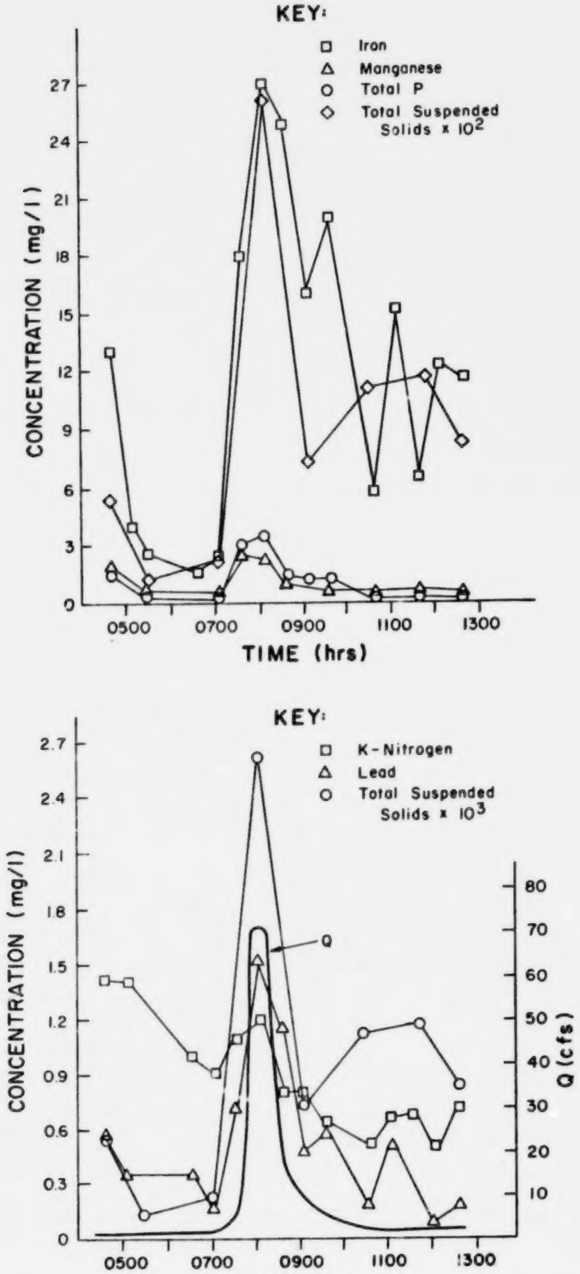


Figure 5. Pollutant variations with Q and time for storm no. 20, date 6/20/72. (Redrawn after Colston (10))

TABLE 2. TOTAL SUSPENDED SOLIDS AND ORGANICS DATA AVERAGED FOR 36 STORM EVENTS (DATA OF COLSTON (10)) AND RATIOS OF ORGANICS TO TOTAL SUSPENDED SOLIDS

Storm Number	Total Suspended Solids (TSS) mg/l		COD mg/l		TOC mg/l		BOD mg/l		COD/TSS	TOC/TSS	BOD/TSS
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ			
1	89	38	25	14			18	14	.28	--	.20
2	274	164	259	62					.95	--	--
3	163	86	111	21	30	7			.68	.18	--
4			171	45	36	7			--	--	--
5	346	272	146	89	35	34	18	13	.42	.10	.05
6	474	249	141	60	25	11	17	12	.30	.05	.04
7	1459	535	195	103	36	41	6	6	.13	.02	<.01
8	1233	949	143	104	33	16			.12	.03	--
9	1754	1194	149	116	24	17	2	.4	.08	.01	<.01
10	572	421	125	96	36	27			.22	.06	--
11	990	733	171	146	36	25			.17	.04	--
12	146	58	82	39	36	10	15	11	.56	.25	.10
13	687	472	176	144	44	30	20	12	.26	.06	.03
14	1087	492	123	73	46	20			.11	.04	--
15	843	429	89	49	36	12	18	9	.11	.04	.02
16	2596	2107	257	190	17	12			.10	.01	--
17	1525	655	150	175	15	8	42	11	.10	.01	.03
18	849	1117	41	7	16	5			.05	.02	--
19	899	576	144	106	41	25	5	3	.16	.05	.01
20	895	789	220	135	39	18	55	14	.25	.04	.06
21	2732	725	271	130	73	30	105	23	.10	.03	.04

TABLE 2 (continued)

Storm Number	Total Suspended Solids (TSS) mg/l		COD mg/l		TOC mg/l		BOD mg/l		COD/TSS	TOC/TSS	BOD/TSS
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ			
22	2332	1090	402	430	165	148	73	10	.17	.07	.03
23	554	290	96	52	26	9	100	5	.17	.05	.18
24	2889	1266	348	198	94	41	80	19	.12	.03	.03
25	--	--	187	79	48	14	16	2	--	--	--
26	3913	2204	184	80	50	18	220	10	.05	.01	.06
27	2522	2434	253	232	51	41	41	24	.10	.02	.02
28	1024	376	140	60	21	11	--	--	.14	.02	--
29	1326	624	142	59	38	16	138	15	.11	.03	.10
30	1340	1100	157	69	44	13	182	60	.12	.03	.14
31*	83	62	132	83	49	15	80	74	1.59	.59	.96
32	777	788	110	77	34	10	--	--	.14	.04	--
33	1246	550	93	28	38	14	49	20	.07	.03	.04
34	1463	923	374	103	105	35	50	12	.26	.07	.03
35	1029	288	289	101	99	19	100	20	.28	.10	.10
36	643	202	92	31	31	14	--	--	.14	.05	--
									*Range: .05 → .95 .95/.05=19	.01 → .25 .25/.01=25	<.01 → .20 .20/.01=20

*Note that storm 31 has not been considered as it may represent erroneous values.

TABLE 3. TOTAL PHOSPHORUS, KJELDAHL NITROGEN, AND FECAL COLIFORMS IN RUNOFF AVERAGED FOR 36 STORM EVENTS (DATA OF COLSTON (10)) AND RATIOS OF EACH TO TOTAL SUSPENDED SOLIDS

Storm Number	Total P mg/l		K-Nitrogen mg/l		Fecal Coliforms #/ml		Total P/ TSS(x10^{-3})	Kjeldahl N/ TSS(x10^{-3})	Fecal Coliforms/ TSS(x10^{-3})
	Avg.	σ	Avg.	σ	Avg.	σ			
1	0.28	0.04	0.91	0.09			3.15	10.2	--
2	0.6	0.21					2.19	--	--
3	1.03	0.35	9.52*	4.01*	203	148	6.32	58.4*	1245
4	0.47	0.13	9.68*	3.57*	398	104	--	--	--
5	1.05	0.86	1.30	0.17	387	246	3.03	3.76	1118
6	0.75	0.31	1.94	0.49	689	111	1.58	4.09	1454
7	1.07	1.05	0.67	0.59	106	68	0.73	0.46	73
8	0.58	0.24	0.55	0.31	74	39	0.47	0.45	60
9	0.5	0.43	0.56	0.46	54	48	0.29	0.32	31
10	0.57	0.26	0.59	0.12	67	28	1.00	1.03	117
11	1.05	0.36	0.76	0.30			1.06	0.77	--
12	0.81	0.22	0.65	0.15	102	63	5.55	4.45	699
13	1.98	4.65	0.6	0.21	137	71	2.88	0.87	199
14	--	--	--	--	--	--	--	--	--
15	--	--	--	--	--	--	--	--	--
16	1.03	0.79	0.67	0.58	143	43	0.40	0.26	55
17	0.56	0.23	0.44	0.2	161	51	0.37	0.29	106
18	0.56	0.29	0.31	0.03	4	2	0.66	0.37	5
19	0.62	0.34	0.82	0.12	442	387	0.69	0.91	492
20	1.09	0.93	0.88	0.29	98	83	1.22	0.98	110
21	1.17	0.55	0.49	0.25	--	--	0.43	0.18	--

TABLE 3 (continued)

Storm Number	Total P mg/l		K-Nitrogen mg/l		Fecal Coliforms #/ml		Total P/ TSS(x10^{-3})	Kjeldahl N/ TSS(x10^{-3})	Fecal Coliforms/ TSS(x10^{-3})
	Avg.	σ	Avg.	σ	Avg.	σ			
22	1.13	0.39	0.64	0.33	175	94	0.48	0.27	75
23	0.44	0.19	0.25	0.08	280	125	0.79	0.45	505
24	1.42	0.37	0.34	0.25	549	363	0.49	0.12	190
25	0.73	0.23	0.35	0.11	--	--	--	--	--
26	--	--	--	--	--	--	--	--	--
27	0.85	0.47	0.48	0.19	172	221	0.34	0.19	68
28	0.36	0.07	0.33	0.14	248	65	0.35	0.32	242
29	0.71	0.21	2.07	1.28	--	--	0.54	1.56	--
30	.71	.24	.43	.13	94	31	0.53	0.32	70
31	0.92	0.62	0.34	0.05	273	168	11.08	4.10	3289
32	--	--	--	--	--	--	--	--	--
33	0.6	0.18	0.2	0.05	44	21	0.48	0.16	35
34	1.54	0.51	0.70	0.13	--	--	1.05	0.48	--
35	--	--	--	--	--	--	--	--	--
36	0.59	0.13	0.48	0.18	258	105	0.92	0.75	401
							Range: 0.29 → 11.08 11.08/.29=38.2	0.12 → 10.2 10.2/.12=85.0	5 → 3289 3289/5=658

*Questionable values.

TABLE 4. METALS CONCENTRATIONS (Ca, Co, Cu, Cr, Fe) AVERAGED FOR 36 STORM EVENTS (DATA OF COLSTON (10)) AND RATIOS OF EACH TO TOTAL SUSPENDED SOLIDS

Storm Number	Calcium mg/l		Cobalt mg/l		Copper mg/l		Chromium mg/l		Iron mg/l		Ca/TSS $(x10^{-3})$	Co/TSS $(x10^{-3})$	Cu/TSS $(x10^{-3})$	Cr/TSS $(x10^{-3})$	Fe/TSS $(x10^{-3})$
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ					
1	--	--	0.36	0.09	0.36	0.10	--	--	4.4	1.3	--	4.04	4.04	--	49.4
2	--	--	0.13	0.03	.14	.03	.31	.07	3.5	.8	--	0.47	0.51	1.13	12.8
3	--	--	0.08	0.04	0.10	0.03	0.33	0.08	3.8	1.7	--	0.49	0.61	2.02	23.3
4	--	--	0.14	0.05	0.13	0.03	0.31	0.06	10.6	2.9	--	--	--	--	--
5	7.0	7.8	--	--	--	--	0.27	0.03	9.1	4.2	20.2	--	--	0.78	26.3
6	2.5	0.5	--	--	--	--	--	--	11.2	4.1	5.27	--	--	--	23.6
7	2.7	1.6	--	--	--	--	0.29	0.09	11.4	4.9	1.85	--	--	0.20	7.81
8	5.5	7.9	--	--	--	--	0.27	0.07	9.9	6.3	4.46	--	--	0.22	8.03
9	4.1	1.3	--	--	--	--	--	--	9.1	4.7	2.34	--	--	--	5.19
10	14.3	14.2	--	--	--	--	--	--	7.8	3.9	25.00	--	--	--	13.6
11	6.2	3.5	--	--	--	--	--	--	16.3	14.5	6.26	--	--	--	16.5
12	24.7	4.9	--	--	--	--	--	--	3.7	1.0	16.9	--	--	--	25.3
13	5.3	1.3	--	--	--	--	--	--	13.6	9.7	36.3	--	--	--	93.2
14	4.2	1.4	--	--	--	--	--	--	9.3	3.5	3.86	--	--	--	8.56
15	5.2	3.7	--	--	--	--	--	--	7.7	4.4	6.17	--	--	--	9.13
16	--	--	--	--	0.15	0.07	0.16	0.09	--	--	--	--	0.06	0.06	--
17	--	--	--	--	0.10	0.00	0.11	0.03	--	--	--	--	0.07	0.07	--
18	--	--	--	--	0.10	0.00	0.10	0.00	--	--	--	--	0.12	0.12	--
19	--	--	0.1	0.0	--	--	--	--	12.9	8.3	--	0.11	--	--	14.4
20	--	--	0.09	0.00	--	--	--	--	12.1	8.2	--	0.10	--	--	13.5
21	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

TABLE 4 (continued)

Storm Number	Calcium mg/1		Cobalt mg/1		Copper mg/1		Chromium mg/1		Iron mg/1		Ca/TSS $(x10^{-3})$	Co/TSS $(x10^{-3})$	Cu/TSS $(x10^{-3})$	Cr/TSS $(x10^{-3})$	Fe/TSS $(x10^{-3})$
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ					
22	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
23	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
24	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
25	2.3	0.8	--	--	0.14	0.02	0.10	0.04	--	--	--	--	--	--	--
26	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
27	--	--	--	--	--	--	--	--	32.8	14.6	--	--	--	--	13.0
28	--	--	--	--	--	--	--	--	19.0	5.6	--	--	--	--	18.6
29	2.1	1.3	--	--	0.12	0.01	0.15	0.04	--	--	1.58	--	0.09	0.11	--
30	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
31	--	--	--	--	--	--	--	--	18.8	9.3	--	--	--	--	227.
32	--	--	--	--	--	--	--	--	19.6	9.4	--	--	--	--	25.2
33	--	--	--	--	0.12	0.02	--	--	--	--	--	--	0.10	--	--
34	--	--	--	--	0.12	0.02	0.16	0.03	--	--	--	--	0.08	0.11	--
35	--	--	--	--	0.13	0	0.11	0	--	--	--	--	0.13	0.11	--
36	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
											Range: 1.58 → 169 169/1.59=107	0.1 → 4.04 4.04/.1=40.4	.06 → 4.04 4.04/.06=67.3	.06 → 2.02 2.02/.06=33.7	5.19 → 227 227/5.19=43.7

TABLE 5. METALS CONCENTRATIONS (Pb, Ni, Mg, Mn, Zn) AVERAGED FOR 36 STORM EVENTS (DATA OF COLSTON (10)) AND RATIOS OF EACH TO TOTAL SUSPENDED SOLIDS

Storm Number	Lead mg/l		Nickel mg/l		Magnesium mg/l		Manganese mg/l		Zinc mg/l		Pb/TSS ($x10^{-3}$)	Ni/TSS ($x10^{-3}$)	Mg/TSS ($x10^{-3}$)	Mn/TSS ($x10^{3}$)	Zn/TSS ($x10^{-3}$)
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ					
1	--	--	--	--	--	--	--	--	0.27	0.27	--	--	--	--	3.03
2	.49	.11	.16	.04	--	--	.59	.10	.19	.04	1.79	0.58	--	2.15	0.69
3	0.43	0.07	0.16	0.03	--	--	0.71	0.26	0.17	0.08	2.64	0.98	--	4.36	1.04
4	0.53	0.09	0.18	0.04	--	--	0.60	0.11	0.32	0.08	--	--	--	--	--
5	0.57	0.24	--	--	11.7	4.7	0.62	0.24	0.31	0.15	1.65	--	33.8	1.79	0.90
6	0.40	0.14	--	--	9.3	2.1	0.63	0.23	0.24	0.10	0.84	--	19.6	1.33	0.51
7	0.42	0.22	--	--	9.1	3.5	0.69	0.37	0.34	0.19	0.29	--	6.24	0.47	0.23
8	0.43	0.16	--	--	10.6	5.4	0.56	0.26	0.27	0.15	0.35	--	8.60	0.45	0.22
9	0.35	0.26	--	--	8.5	2.2	0.52	0.16	0.68	0.91	0.20	--	4.85	0.30	0.39
10	0.47	0.30	--	--	11.2	3.2	0.64	0.14	0.58	0.29	0.82	--	19.6	1.12	1.01
11	0.57	0.80	--	--	15.2	4.0	0.89	0.44	0.96	0.88	0.58	--	15.4	0.90	0.97
12	0.26	0.12	--	--	9.4	1.6	0.44	0.06	0.21	0.06	1.78	--	64.4	3.01	1.44
13	0.38	0.29	--	--	3.9	.3	0.51	0.29	0.36	0.29	0.55	--	5.68	0.74	0.52
14	0.29	0.14	--	--	8.6	2.4	0.43	0.22	0.25	0.10	0.27	--	7.91	0.40	0.23
15	0.23	0.15	--	--	7.7	1.1	0.33	0.18	0.21	0.11	0.27	--	9.13	0.39	0.25
16	0.45	0.46	--	--	--	--	--	--	0.50	0.37	0.17	--	--	--	0.19
17	0.23	0.17	--	--	--	--	--	--	0.25	0.16	0.15	--	--	--	0.16
18	0.10	0.00	--	--	--	--	--	--	0.10	0.01	0.12	--	--	--	0.12
19	0.47	0.54	0.1	0.0	--	--	1.07	0.76	--	--	0.52	0.11	--	1.19	--
20	0.49	0.38	0.09	0.00	--	--	0.95	0.63	--	--	0.55	0.10	--	1.06	--
21	--	--	--	--	13.4	1.4	2.01	0.38	0.83	0.37	--	--	4.90	0.74	0.30

TABLE 5 (continued)

Storm Number	Lead mg/l		Nickel mg/l		Magnesium mg/l		Manganese mg/l		Zinc mg/l		Pb/TSS ($x10^{-3}$)	Ni/TSS ($x10^{-3}$)	Mg/TSS ($x10^{-3}$)	Mn/TSS ($x10^{-3}$)	Zn/TSS ($x10^{-3}$)
	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ	Avg.	σ					
22	--	--	--	--	12.4	1.2	0.71	0.17	0.42	0.06	--	--	5.32	0.30	0.18
23	--	--	--	--	4.7	0.8	0.40	0.25	0.22	0.10	--	--	8.48	0.72	0.40
24	--	--	--	--	11.9	1.6	1.67	0.24	0.53	0.19	--	--	4.12	0.58	0.18
25	--	--	--	--	7.2	2.2	--	--	0.34	0.11	--	--	--	--	--
26	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
27	0.79	0.75	--	--	--	--	1.32	0.91	--	--	0.31	--	--	0.52	--
28	0.26	0.09	--	--	--	--	0.60	0.16	--	--	0.25	--	--	0.59	--
29	--	--	--	--	12.6	4.4	--	--	--	--	--	--	9.50	--	--
30	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
31	0.20	0.26	--	--	--	--	--	--	0.33	0.23	2.41	--	--	--	3.98
32	0.28	0.12	--	--	--	--	--	--	0.28	0.14	0.36	--	--	--	0.36
33	--	--	--	--	--	--	0.44	0.08	0.28	0.05	--	--	--	0.35	0.22
34	1.19	0.32	--	--	15.5	2.5	--	--	--	--	0.81	--	10.6	--	--
35	0.69	0.14	--	--	--	--	--	--	--	--	0.67	--	--	--	--
36	0.24	0.08	--	--	--	--	--	--	--	--	0.37	--	--	--	--
											Range: 0.12 → 2.64 2.64/.12=22.0	0.10 → 0.98 .98/.1=9.8	4.12 → 64.4 64.4/4.12=15.6	0.30 → 4.36 4.36/.3=14.5	0.12 → 3.98 3.98/.12=33.2

be lower than for ratios of individual observations. Second, a similar moderating effect on variability is probably present in the data because of the large drainage area involved and the distance over which the runoff traveled before being sampled. All of this strongly suggests that the ratios would be substantially more variable if computed for individual samples taken at upshed locations. Finally, it should be noted that the data within each storm are likely to be autocorrelated. This means that the data, not being independent of time, may exhibit variable potency factors from storm to storm because of the sampling schedule. Thus, in examining Colston's (10) data it would have been much better had individual observations been available rather than storm means.

Donigian and Crawford (6), in a study specifically designed to evaluate the feasibility of simulating nutrient washoff using potency factors and the NPS model, also looked at the Third Fork Creek (Durham, North Carolina) Watershed (see Colston (10) and preceding discussion). Donigian and Crawford simulated sediment loss, TKN, total phosphorus, and iron during five storms. Their results showed that for the specific watershed, total phosphorus and iron concentrations in the runoff could generally be simulated with reasonable accuracy. TKN, on the other hand, proved difficult to predict, this being attributable (according to the authors) to subsurface carriage of TKN and miscalibration of the potency factors based upon available data.

Non-Urban Watershed Studies

Donigian and Crawford (6) have applied the NPS model to data developed on agricultural test plots located in Watkinsville, Georgia, and at East Lansing, Michigan. Characteristics of these test plots have been described in detail by Smith *et al*. (11) (Watkinsville) and by Ellis *et al*. (12) (East Lansing). In general, the test plots considered by Donigian and Crawford are the P-2 plot at Watkinsville and the P-6 plot at East Lansing. According to Smith *et al*. (11), the P-2 plot is an area of 1.29 hectares representing common Piedmont forms and management practices. Soil types are primarily Cecil sandy loam with some Cecil sandy clay loam and loam. No soil or water conservation measures were applied in the period during which the data used by Donigian and Crawford were collected. Watershed P-6 is 0.8 hectares in area and consists primarily of Spinks loamy fine sand with some Hillsdale fine sandy loam (12). During the period considered by Donigian and Crawford, plot P-2 was in corn and plot P-6 was in soybeans.

Results of simulation of storm events on the two watersheds with the NPS model are characterized as reasonable (6), with discrepancies between observed and simulated concentrations of total phosphorus, phosphate, nitrate, ammonia, and total nitrogen being attributed primarily to the small size of the test watersheds, the short duration thunder storms, and the inability of the NPS model to accommodate tillage practices. Donigian and Crawford (6) conclude that potency factors appear useful for simulating total nutrients. Total phosphorus concentration closely mirrored those of sediment. Total nitrogen also could be adequately represented through potency factors, although not as well as total P; whereas simulating ammonia, nitrate, and phosphate as a function of sediment load was increasingly tenuous.

The issue of effects of tillage, as noted by Donigian and Crawford (6), was addressed some years earlier, at least for N and P in transported sediment, by Romkens et al. (13). In the study, which was conducted at Bedford, Indiana, five different tillage-planting systems were used with simulated rainfall providing runoff events. It was noted, first, that nitrogen and phosphorus concentrations in the transported sediment were related to runoff sediment (both expressed as mass per unit area of watershed) although the relationship was not necessarily linear. Second, impacts of tillage methods on nitrogen and phosphorus losses were found to be significant, and the tillage method influenced the propensity of fine clay particles to wash off. Finally, correlations between sediment nutrient (N and P) concentrations and sediment clay concentrations were very high (.98 and .96, respectively).

Burwell et al. (14) have noted that, although large amounts of nutrients may be associated with runoff sediment, subsurface transport can also account for substantial amounts of soluble nitrogen and phosphorus. This was also cited by Donigian and Crawford (6) as one factor affecting their results in applying the NPS model, particularly with respect to the poor nitrate and phosphate simulations. Burwell et al. (15) have also observed significant impacts of soil cover and seasonal periods on nutrient transport thus further emphasizing the possible effects of land use and management practices on potency factors.

SECTION 4

METHODOLOGY

The study consisted initially of compilation of data bases representing agricultural, silvicultural, urban, and rural watersheds, and subjecting the data to regression analysis. Subsequent analysis was geared toward determining two things. First, the nature and quality (in terms of a simple mathematical expression) of the relationship between suspended solids content of runoff and other water quality characteristics was determined. Second, the degree to which runoff water quality may be predicted using information on storm characteristics and other variables in addition to suspended sediment loads was evaluated.

GENERAL ANALYSIS PROCEDURE

Figure 6 shows the general procedure by which data were subjected to analysis. The first step in the process consisted of fitting a simple linear model of the form:

$$C = b_0 + b_1 S \tag{3}$$

where C = predicted concentration of water quality constituent

S = concentration of suspended solids

b_0, b_1 = fitted coefficients

The program for fitting the simple linear model provided several useful kinds of output including a scatter plot of the data. Using the plots, it was possible to determine whether there was any obvious nonlinearity in the data as well as to subjectively determine whether regression assumptions were met. Where it was apparent that transforming the data might improve the quality of relationships, appropriate transformations were used, and the plotting/linear regression package was again applied. After it was clear for each data set that further transformation was likely to be of little or no benefit in terms of improvement of fit, processing moved on to multiple regression. In performing the multiple regressions, other possible transformations were examined. Unless significantly better relationships were obtained by using other transformations, the transformations used for simple regression were adopted in multiple regression.

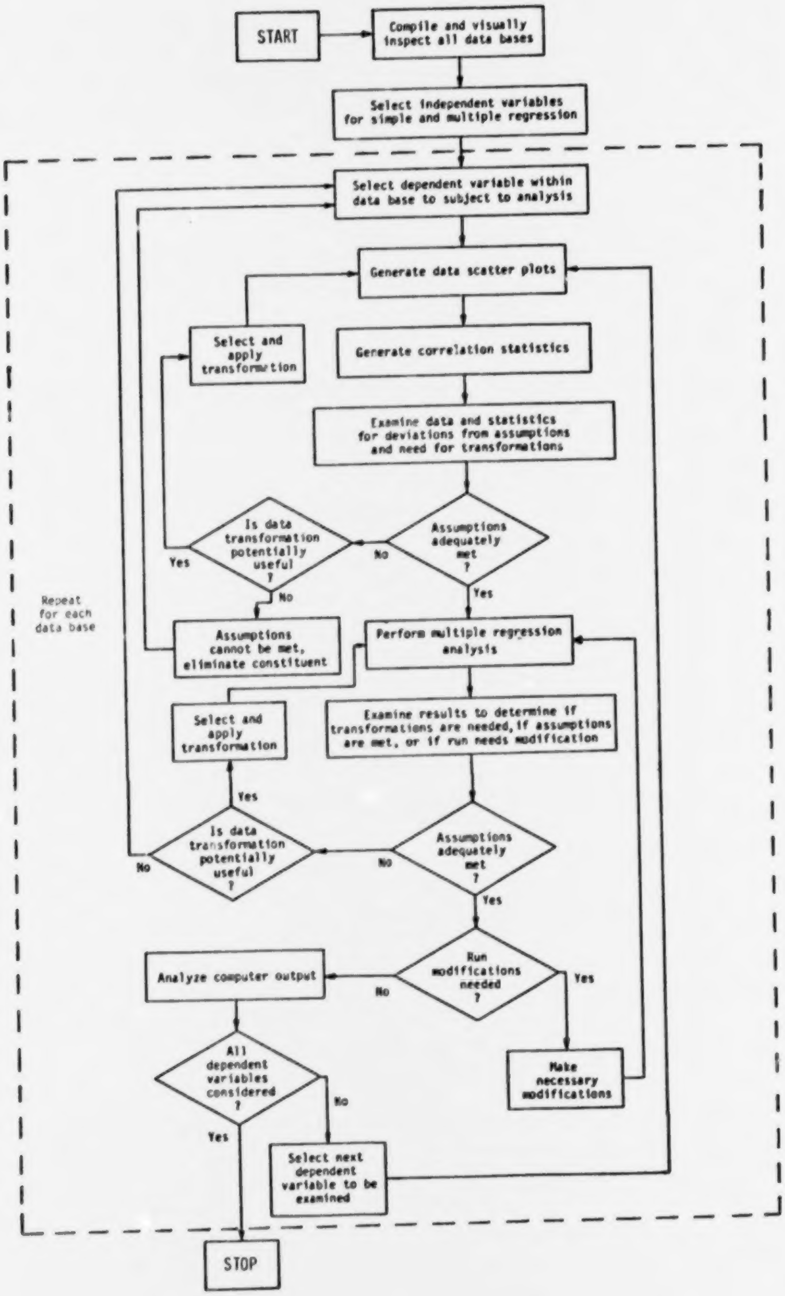

START
Compile and visually inspect all data bases
Select independent variables for simple and multiple regression
Select dependent variable within data base to subject to analysis
Generate data scatter plots
Select and apply transformation
Generate correlation statistics
Examine data and statistics for deviations from assumptions and need for transformations
Is data transformation potentially useful ?
Yes
No
Assumptions adequately met ?
No
Yes
Repeat for each data base
Assumptions cannot be met, eliminate constituent
Perform multiple regression analysis
Select and apply transformation
Examine results to determine if transformations are needed, if assumptions are met, or if run needs modification
Yes
Is data transformation potentially useful ?
No
Assumptions adequately met ?
No
Yes
Analyze computer output
Run modifications needed ?
No
Yes
All dependent variables considered ?
No
Yes
Make necessary modifications
Select next dependent variable to be examined
STOP

Figure 6. General data analysis procedure.

Because the details of statistical techniques in general and regression, in particular, are not universally understood, some general theoretical considerations are presented in Appendix A. The reader wishing to review the materials, or requiring clarification of some point should refer to this Appendix.

Analytical Tools

Figure 6 shows the general procedure by which analyses were performed in this study. The two distinct computer programs that were applied are described below.

Simple Regression--

The first program provided graphical output and certain basic statistics using the simple linear model:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{4}$$

and

$$\hat{Y}_i = b_0 + b_1 X_i \tag{5}$$

where Y_i = ith observed value of the dependent variable Y

X_i = ith observed value of the independent variable X

b_0 = Y intercept of the fitted line, an estimator of β_0

b_1 = slope of the fitted line, an estimator of β_1

β_0 = Y intercept of the true relationship

β_1 = slope of the true relationship

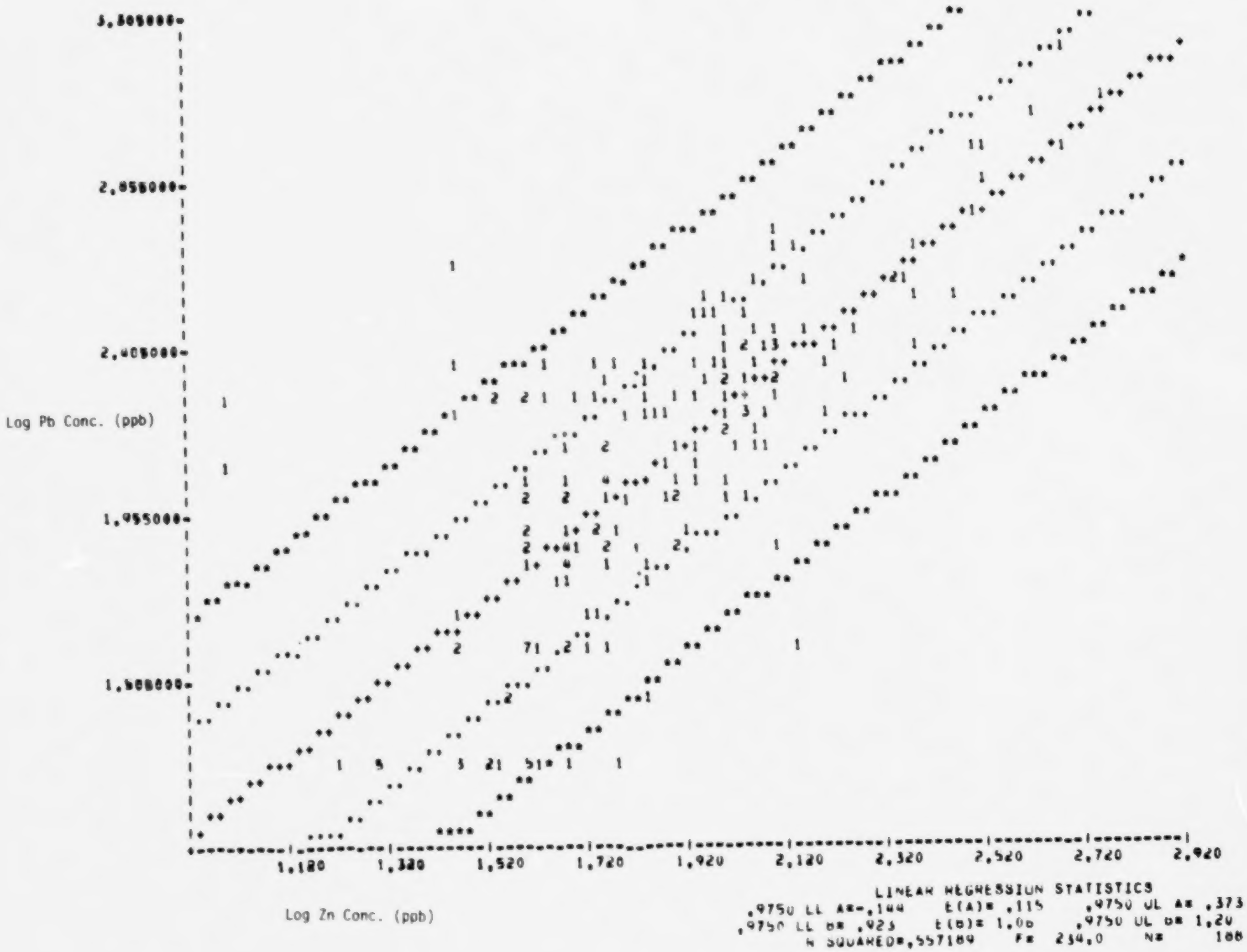
ε_i = error term correcting the value of the dependent variable, for deviations from that described by $\beta_0 + \beta_1 X_i$

$\hat{Y}_i$ = ith predicted value of the dependent variable

The output of the simple regression program appears as shown in Figure 7. In the figure, and considering only the two variables ($\log_{10}$ lead concentration versus $\log_{10}$ zinc concentration), the following are provided.

- Linear and nonlinear relationships (through appropriate data transformations) of each bivariate population. The fitted line is represented by "+" symbols, the number of observations per plot coordinate by plotted numerals.

- Confidence limits for prediction of the population mean and individual values of the dependent (ordinate) variable for a given value of the independent (abscissa) variable. The 95%



```
                 LINEAR REGRESSION STATISTICS
  ,9750 LL A=-,144    E(A)= ,115    ,9750 UL A= ,373
 ,9750 LL B= ,923    E(B)= 1,06    ,9750 UL B= 1,20
       R SQUARED=,557184    F= 234,0    N=     108
```

Figure 7. Sample plot showing output on a simple linear regression program used in this study.

confidence limits are shown on the plot with the band for prediction of the population mean shown as "•" symbols and the band for prediction of individual dependent variable values shown as "*" symbols.

- Estimates for the slope and intercept of the least squares model (designated E (B) and E (A), respectively, in Figure 7).

- Confidence limits for the slope and intercept. In Figure 7, these are designated as LL B, UL B, LL A, UL A for the lower and upper limits of the slope and intercept, respectively.

- The r-squared and F values for the fitted relationship. The former describes the proportion of the variance shared by the ordinate and abscissa variables (given a linear relationship), while the latter, given that regression assumptions are met, indicates whether the value of r-squared is significant at the selected α level. N is the number of observations.

Multiple Regression--

The multiple regression program uses a stepwise algorithm in fitting an equation of the form:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + . . . + b_mX_m \qquad (5)$$

where $\hat{Y}$ = the predicted value of the dependent variable

X_m = the observed values of the independent variables

b_m = the fitted coefficients

In this type of program, the independent variables are introduced into the regression one at a time (one per step), and for each step, all statistics are generated. The criterion for selecting variables is the increase in R^2 (squared multiple correlation coefficient). Thus the first step introduces the independent variable having the highest correlation with the dependent variable. The second step introduces into the regression that independent variable which, in conjunction with the first variable, gives the greatest R^2. The algorithm can also remove a variable which, although introduced in an earlier step, no longer contributes significantly to R^2.

Table 6 shows an example multiple regression. Specifically, a typical output table from the regression package used in this study is shown. Table 7 provides an explanation of entries in Table 6.

TABLE 6. REGRESSION OF LOG $[NO_3 + NO_2]$ (SEATTLE DATA)*

STEP NO. 6
VARIABLE ENTERING 1
FLEVEL 6.034608
STANDARD ERROR OF Y
CONSTANT 2.459798

VARIABLE	COEFFICIENT	STD ERROR OF COEFF	T-STATISTIC
X 1	-.55947	.22775	-2.4565
X 5	-.09976	.02747	-3.6314
X 6	-1.67359	.36013	-4.6471
X 9	.20843	.02662	7.8283
X 13	.00416	.00034	12.2785
X 15	-.13748	.04574	-3.0054

PREDICTED VS. ACTUAL RESULTS

OB. NO.	ACTUAL	PREDICTED	DEVIATION	PERCENT DEV.
1	-.2076	-.2886	.1055	17.01
2	-.2147	-.2203	.0079	1.29
3	-.1675	-.1562	-.0179	-2.63
4	-.1739	-.1554	-.0292	-4.36
5	-.1739	-.2095	.0527	7.87
6	-.2676	-.2724	.0060	1.10
7	-.3468	-.2939	-.0583	-12.95
8	-.2596	-.2112	-.0650	-11.81
9	.4472	.0083	1.7808	63.60
10	.3424	.0532	1.0696	48.62
181	-.7447	-.7084	-.0157	-8.72
182	-.6198	-.6144	-.0030	-1.24
183	-.6383	-.6207	-.0095	-4.13
184	-.6778	-.6259	-.0267	-12.70
185	-.1487	.0031	-.2973	-41.87
186	.0128	.0032	.0225	2.18
187	-.1612	-.2354	.1085	15.72
188	.1461	.2420	-.3458	-24.70

NO.	INDEPENDENT VARIABLE CODES
1	(LOG OF) DAYS SINCE 1/1/73
2	MONTH
3	HOURS SINCE STORM START

TABLE 6 (continued)

```
                    4     (LOG OF) DRY DAYS BEFORE STORM
                    5     (LOG OF) SUSPENDED SOLIDS, MG/L
                    6     (LOG OF) CD, UG/L
                    7     (LOG OF) PB, UG/L
                    8     (LOG OF) ZN, UG/L
                    9     (LOG OF) NH4, MG/L-N
                   10     (LOG OF) ORGANIC N, MG/L-N
                   11     (LOG OF) TOTAL P, MG/L-P
                   12     (LOG OF) OPO4-P, MG/L-P
                   13              CONDUCTIVITY, U-MHOS/CM
                   14     (LOG OF) TURBIDITY, JTU
                   15     (LOG OF) FLOW, CFS
                   16              STORM RAINFALL, INCHES

R-SQUARED FOR THE REGRESSION LINE IS  .753564

CORRECTED R-SQUARED IS  .745395

F FOR THE REGRESSION IS    92.25    F CORRECTED FOR AN EQUIVALENT OF
    188 OBSERVS.=  92.25

AUTOREGRESSION COEFFICIENTS FOR THE FIRST TEN DIFFERENCES ARE-
      .111693    .093102    .077329    .010425    -.003380    .000156    .001426
      .000575    .002910   -.001002
```

*Transcribed from computer output, observations 11 through 180 not shown.

TABLE 7. SIGNIFICANCE OF ENTRIES IN TABLE 6

Item	Interpretation
Step No. 6	The program used provides stepwise multiple regression. During each step, either a new variable is added or a previously included one is removed. Table 6 contains the sixth step in the analysis.
Variable entering 1	During this step, variable 1 has been introduced as a predictor variable.
F level	This is the value of F for the current regression step. It reveals whether the step has significantly improved prediction over the preceding step.
Standard error of Y	This is the standard error of the predicted values, also called the "standard error of estimate." It may be used in computing a confidence interval for each predicted value. Although the latter, also called the "standard error of forecast," is not shown in Table 6, it is provided in later runs. The standard error of forecast is not of much use here, where the main objective is explication. It is important, however, where the regression equation is to be used as a predictive model.
Constant	The predictive model is of the general form $\hat{Y} = b_0 + b_1X_1 + b_2X_2 + . . . + b_nX_n$ The constant is b_0, the intercept for the regression equation.
Variable	The independent variable(s) used in the model. In this case, variables 1, 5, 6, 9, 13, and 15 have been used.
Coefficient	These are the "b" values in the regression equation.

TABLE 7 (continued)

Item	Interpretation
Standard error of coefficient	The standard errors of the coefficients are printed out here. These are used in computing the t-statistic.
t-statistic	This is the ratio of the estimated coefficient to the standard error of the coefficient. In general, where the absolute value exceeds 2, the probability that the coefficient is non-zero is at least $p = .95$. Thus, where $\lvert t \rvert \geq 2$, use of the variable is justified on a basis of statistically non-zero coefficient. If $\lvert t \rvert < 2$, there is no justification in saying that the variable predicts $NO_3 + NO_2$ for any reason other than chance (at $p = .95$).
Observation number	The number of the observation.
Actual	Observed concentration of $NO_2 + NO_3$ ($\log_{10}$ transformation). The first observation was $10^{-.2076} = .620$ ppm.
Predicted	Logged concentration of nitrate as predicted by the model $10^{-.2886} = .515$.
Deviation	Untransformed difference between observed and predicted, $= .620 - .515 = .105$.
Percent deviation	The deviation as percent of the observed untransformed value $(.105/.620) \times 100\% = 17\%$.
Independent variable codes	Shows precise identity of each variable used in regression.
R-squared for the regression	Proportion ($0 < R^2 < 1$) of the variance of the dependent variable ($\log_{10} [NO_3 + NO_2]$ in this case) accounted for by the independent variables (the regression model).

TABLE 7 (continued)

Item	Interpretation
Corrected R-squared	R-squared adjusted for the number of variables used in the model. The more variables used, the more R-squared is likely to be inflated, and the greater the correction needed. This value is a more accurate reflection of the R^2 likely to be obtained if a new data set were collected under identical circumstances.
F for this regression	F statistic showing whether or not R^2 for the model is significant. That is, it shows how likely it is that the results have occurred by chance alone.
Autoregression coefficients for the first ten differences	This statistic examines the residuals (deviations) to determine whether there is reason to suspect a lack of independence of observations or residuals (a regression assumption).
Von Neumann ratio for first ten differences	Corroborative statistic for the autoregression coefficient. Printed out in other tables but not shown in Table 6.
F corrected · · · observs. =	F-statistic corrected for the number of equivalent independent observations (here equal to 188) computed from the number of data points (in this case 188) and the autoregression coefficients.

Data Bases

Data for this study represent six different geographical areas and ten discrete watersheds. Five of the watersheds are agricultural, one is silvicultural, one is rural, and three are urban. Table 8 provides a summary description of the data sets used. Appendix B describes water quality analysis methods as described in available reports for the data bases.

Watkinsville--

The Watkinsville data sets were derived from agricultural test plots maintained jointly by the U.S. EPA Environmental Research Laboratory, Athens, Georgia, and the Agricultural Chemical Transport and Modeling Unit of the Southern Piedmont Conservation Research Center, United States Department of Agriculture. The locations and configurations of the test plots are as shown in Figures 8, 9, and 10, and are described in detail in Smith *et al*. (11). Smith and coworkers also describe the experimental procedure and results of several years of runoff sampling and analyses from the test watersheds.

All test plots were constructed to have a drainage channel consisting of a stainless steel flume. Continuous runoff sampling for sediment and pollutants was provided by use of a motorized sampling slot traversing back and forth through the discharge. The slot was tapered so that a larger proportion of low flows than high flows would be collected. Stationary slots below the flume subdivided the sample, and a sequential sampler delivered samples to a refrigerated compartment.

Data from test plot P-01 consisted of runoff concentrations of the pesticides trifluralin, paraquat, and diphenamid (1972, 1973); paraquat and diphenamid (1974); and propazine and paraquat (1975). The data also included rainfall and runoff information with measurements of flume stage height, sediment concentration, and sampling time provided. These data were also provided for plots P-03 and P-04.

Test plot P-02 data were not used in the present study, and this plot will not be discussed here. Test plot P-03 was used for runoff pesticide data with analyses of trifluralin, paraquat, and diphenamid (1972, 1973); and paraquat and diphenamid (1974) being available.

Test plot P-04 was used to develop pesticide data (paraquat and atrazine, 1973 and 1974; and atrazine, cyanizine, paraquat, and 2,4-D, 1975) and nutrient runoff data as well. Nutrient data generated include nitrate, ammonia, TKN (total kjeldahl nitrogen), phosphate, available phosphorus, total phosphorus, and chloride. With respect to chloride, phosphate, and nitrate, only aqueous phase concentrations were measured. Only sediment concentrations of available phosphorus were provided while the remaining constituents were measured both in the sediment and in the aqueous phase.

Storms were monitored from the time of cropping for varying durations on the order of 6 months to 1 year, depending on plot and stage in the study. Within storms, the number of observations varied with up to about 35 taken in some individual storms.

TABLE 8. SUMMARIES FOR DATA BASES USED IN THIS STUDY

Data Base	Location	Type	Area (ha)	Number of Constituents Examined	Number of Observations Used (range)	Approximate Number of Storms**	Period of Record Used in the Study
Watkinsville	Oconee County, Georgia	Agricultural test plots					
Test plot P-1			2.7	2	134-260		7/72-5/74
Test plot P-3			1.26	2	85-123		7/72-10/75
Test plot P-4			1.4	6	21-108	13	5/74-5/75
Buffalo Bill	Near Eldridge, Scott County, Iowa	Agricultural Watershed	1,417	10	42-89	7	8/73-12/73
Michigan State University Farms, Test plot P-6	East Lansing, Ingham County, Michigan	Agricultural test plot	0.8	5	73-148	24	1/75-4/75
Redwood National Park	Humboldt County, California	Silvicultural	*	6	7-17		9/73-9/75
Seattle	Washington						
Viewridge 1		High density, old residential, urban, storm sewer	255	8		30	2/73-9/73, 10/74-12/75
South Seattle		Industrial, urban, storm sewer	11.1	8	184-199	31	5/73-9/73, 10/74-12/75
Southcenter		New shopping center, urban, storm sewer	9.8	8	188-199	30	2/73-9/73, 10/74-12/75
Honey Creek	Melmore, Ohio	Predominantly rural	38,600	11	123-636		1/76-1/77

*Several stations considered, precise drainage areas not available

**Storms having data in the data base for runoff water quality and used in analysis. Only approximate because some storms merged and the data were often unclear on this issue. In some cases (Michigan) some storms were snow storms.

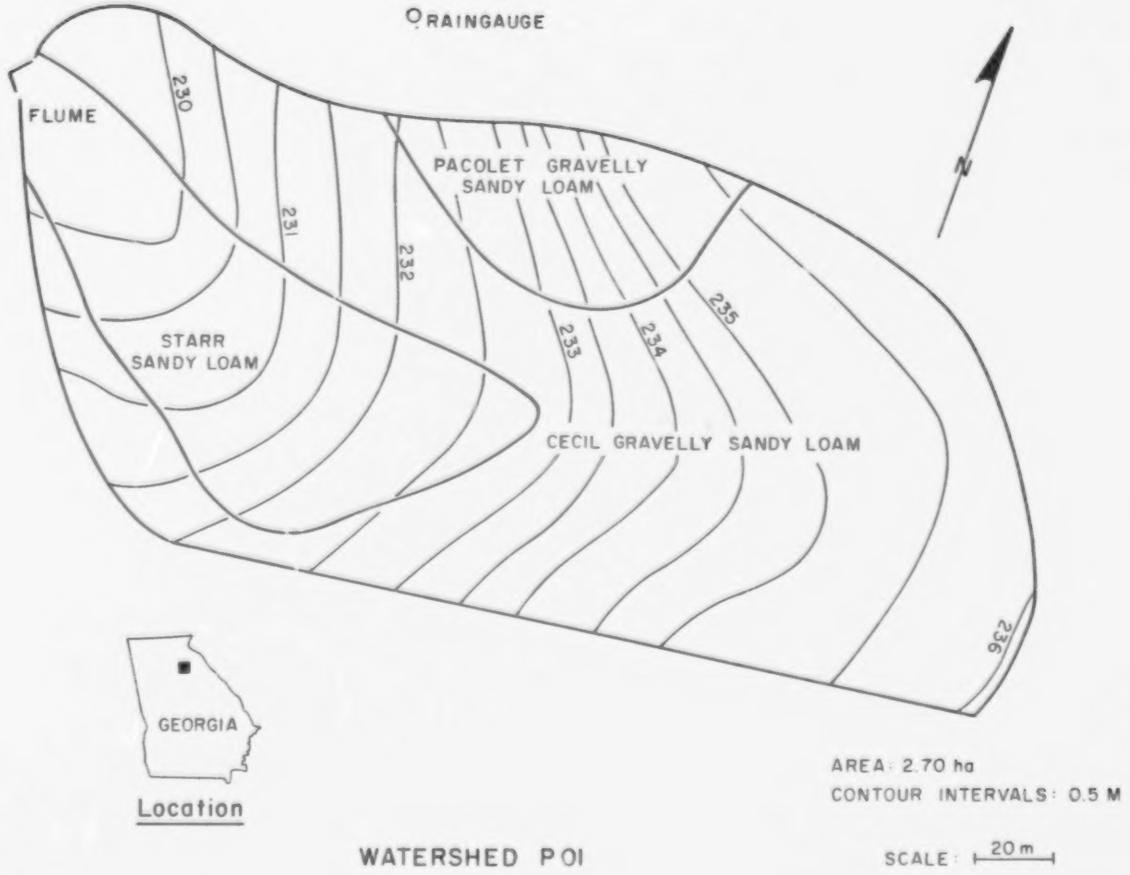

RAINGAUGE
FLUME
230
231
232
233
234
235
236
PACOLET GRAVELLY SANDY LOAM
STARR SANDY LOAM
CECIL GRAVELLY SANDY LOAM
N
GEORGIA
Location
WATERSHED P01
AREA: 2.70 ha
CONTOUR INTERVALS: 0.5 M
SCALE: 20 m

Figure 8. Soils and topography, watershed P-01. (Redrawn after Smith et al. (11))

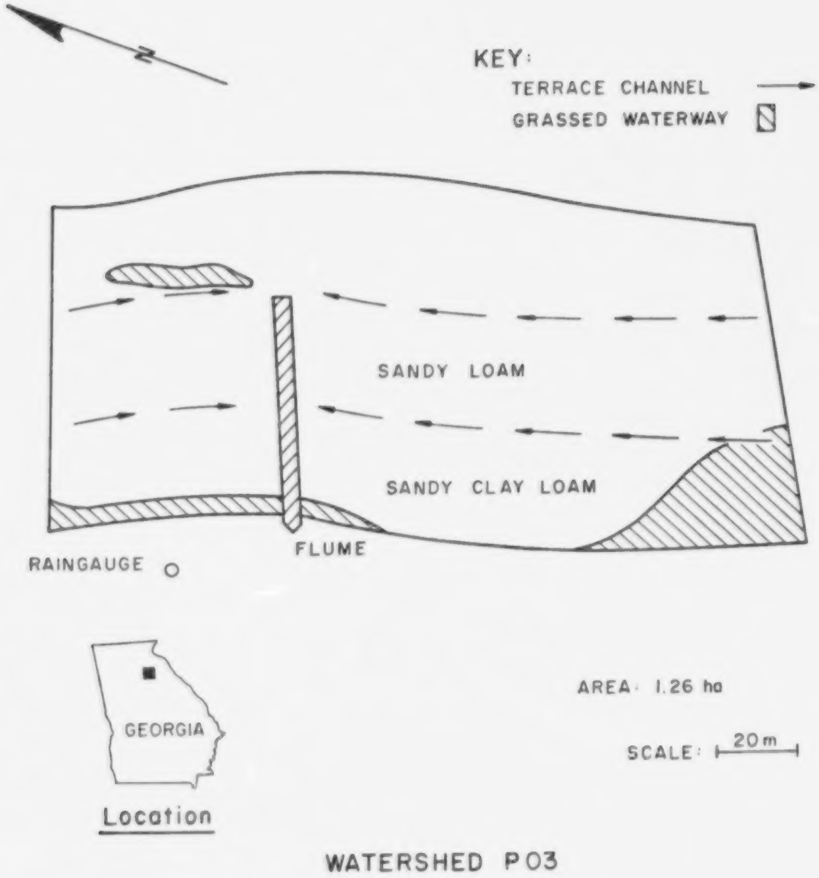
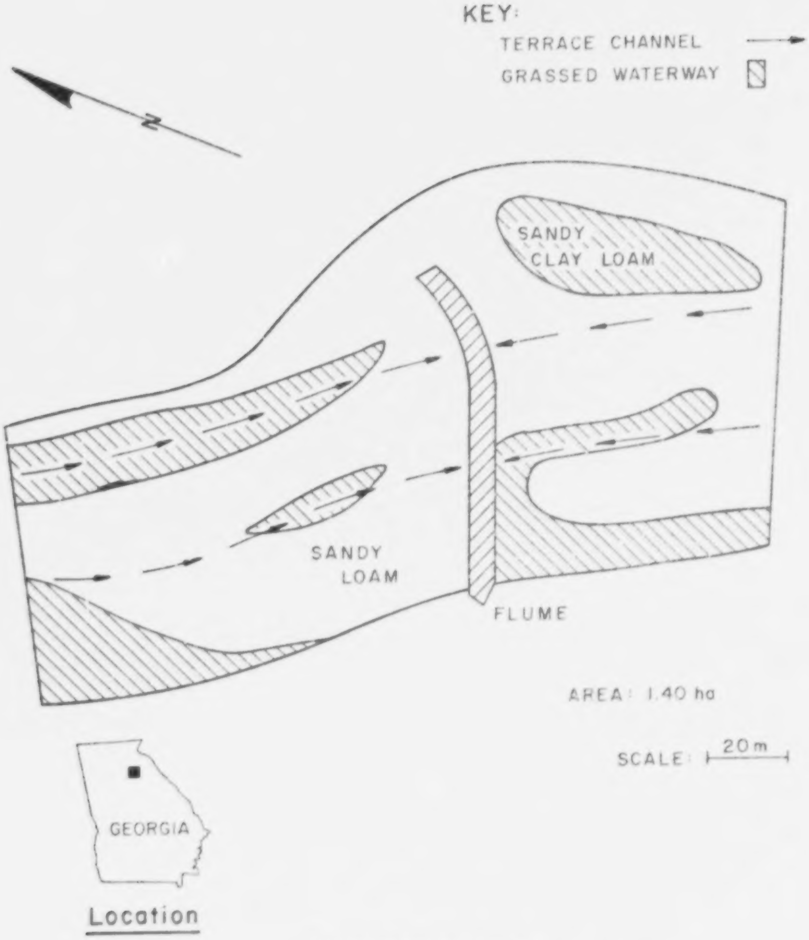

N
KEY:
TERRACE CHANNEL
GRASSED WATERWAY
SANDY LOAM
SANDY CLAY LOAM
RAINGAUGE
FLUME
GEORGIA
Location
AREA: 1.26 ha
SCALE: 20 m
WATERSHED P03

Figure 9. Soils and terrace configurations, watershed P-03. (Redrawn after Smith *et al*. (11))


KEY:
TERRACE CHANNEL
GRASSED WATERWAY
N
SANDY CLAY LOAM
SANDY LOAM
FLUME
AREA: 1.40 ha
SCALE: 20 m
GEORGIA
Location

WATERSHED P 04

Figure 10. Soils and terrace configurations, watershed P-04. (Redrawn after Smith *et al.* (11))

Buffalo Bill Watershed--

The Buffalo Bill Watershed is a 1,417 hectare agricultural basin located in eastern Iowa. During the period in which the data used in this study were compiled, it was primarily in corn, beans, and pasturage, with very small areas in hay and oats. The study from which the data were drawn was conducted by the Iowa State Hygienic Laboratory (associated with the University of Iowa, Iowa City) during the last five months of 1973.

Figure 11 shows the configuration and location of the drainage basin. The nine sampling stations shown are the locations where runoff was sampled. These were pooled for purposes of statistical analysis because of the high quality of the data and relatively few observations available at each station.

Table 9 shows storm data for the Buffalo Bill data base (16). As shown, there were seven storms, with the storms of August 13 and October 10 being the most substantial.

Runoff quality variables included within the data base are as follows:

- Dissolved Oxygen (DO)
- Biochemical Oxygen Demand (BOD)
- Fecal Coliforms
- Total Kjeldahl Nitrogen (TKN)
- Ammonia N
- Nitrite N
- Nitrate N
- Phosphate
- pH
- Suspended Solids (SS)
- DDT
- DDE
- Dieldrin
- Aldrin
- Heptachlor
- Heptachlor epoxide
- Gammachlordane
- Organophosphate
- Turbidity
- Dissolved Solids

Data were generally available for the non-pesticide parameters and for DDT, DDE, and dieldrin, but not for the remaining pesticides. Accordingly, the data on these pesticides (DDT, DDE, dieldrin) and BOD, fecal coliforms, TKN, NH_4, NO_2, NO_3, and phosphate were subjected to analysis. In their discussion of the study, Morris and Johnson (16) do not describe the way in which phosphate, TKN, NH_4, and BOD were measured. Accordingly, it is not known whether these represent dissolved, suspended, or total values. Nor is BOD specified as representing five day, ultimate, or other basis.

Michigan State University Farms--

Data were drawn from a 27-month study of nutrient and pesticide losses from great lakes watersheds, conducted by members of the Departments of Crop and Soil Sciences and of Entomology of the Michigan State University at East

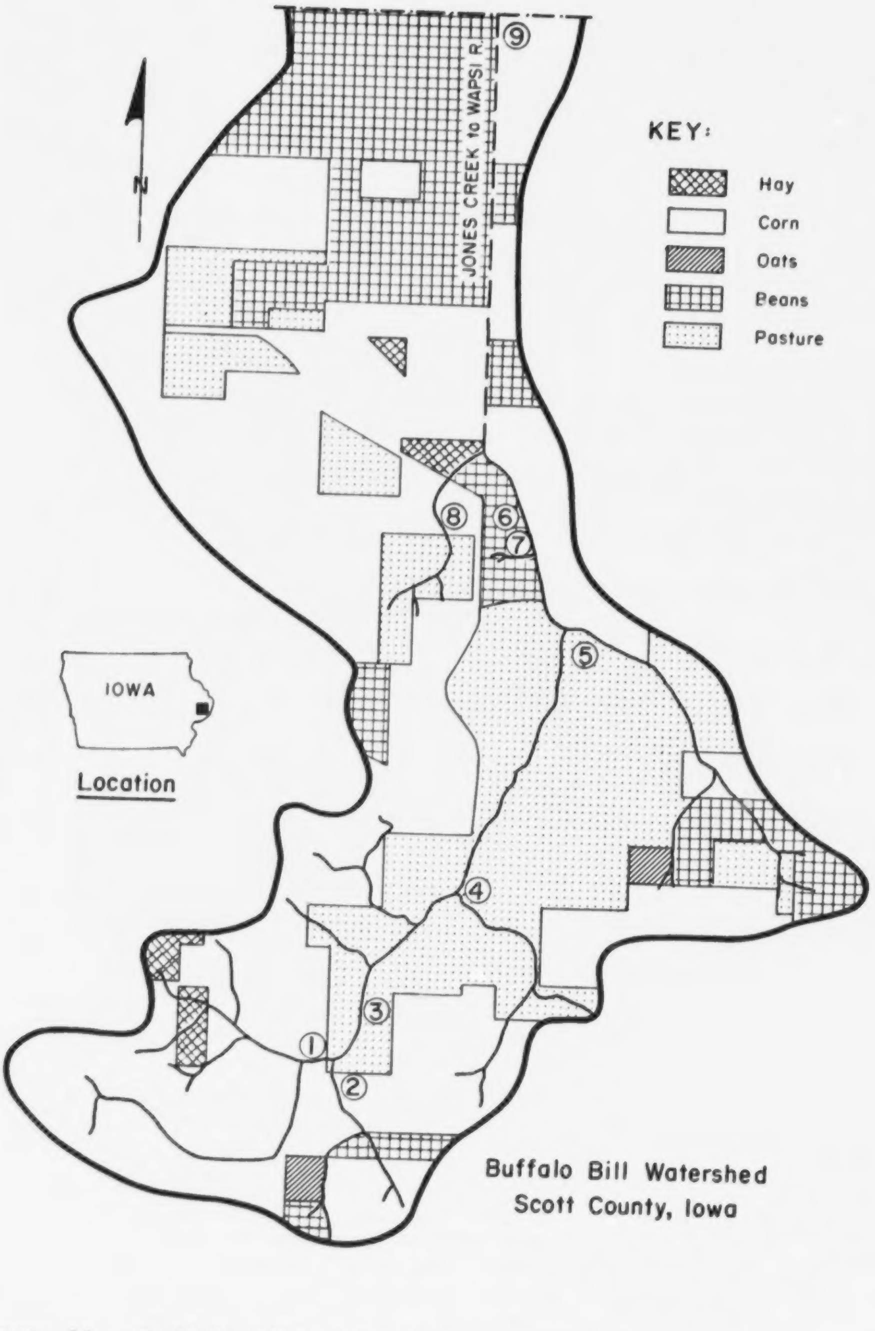

JONES CREEK to WAPSI R.
N
KEY:
Hay
Corn
Oats
Beans
Pasture
IOWA
Location
⑨
⑧
⑥
⑦
⑤
④
③
①
②
Buffalo Bill Watershed
Scott County, Iowa

Figure 11. Buffalo Bill Watershed configuration and location. (Redrawn after Morris and Johnson (16))

TABLE 9. BUFFALO BILL WATERSHED FLOW AND RAINFALL DATA (16)

Date	Time	Estimated Flow (cu ft/sec)	Rainfall (inches)
8/13/73	6:22 pm	2100	2.9
8/13/73	7:10 pm	1300	0.1
8/13/73	10:15 pm	200	0.1
9/04/73	1:00 pm	70	0.5
9/17/73	1:10 pm	50	0.85
9/26/73	8:00 pm	40	0.15
9/26/73	11:20 pm	110	0.7
10/10/73	6:15 pm	380	0.65
10/10/73	8:55 pm	70	0.1
10/29/73	11:15 pm	5	0.1
12/04/73	12:25 pm	45	1.5
12/04/73	3:15 pm	35	0.15
12/12/73	1:45 pm	3	0.0

Data are representative of Station Number 9
(see Figure 11).

Lansing. The study was supported by the EPA Environmental Research Laboratory, Athens, and was a parallel study, both in time (5/73-8/75) and in the sense that it was in support of the Watkinsville study. Ellis *et al.* (12), (17) describe the study in detail in their project reports.

The Michigan study employed two watersheds of 0.8 hectare (east watershed, designated P-06) and of 0.55 hectare (west watershed, designated P-07). Figure 12 shows the location and configuration of the watersheds. During the 10-year period prior to the study, they were in corn, but during the study period, the watersheds were in soybeans.

Pollutants measured in the runoff from the watersheds included paraquat, trifluralin, diphenamid, atrazine, nitrate, total kjeldahl nitrogen, ammonia, total phosphorus, available phosphorus, phosphate, and chloride. With the exception of nitrate, phosphate, and chloride, which were measured only in the water fraction, and suspended solids concentrations of available phosphorus, all pollutants were measured in both the water fraction and in the suspended solids fraction.

Redwood National Park Studies--

The United States Geological Survey (USGS) has conducted a study of sedimentation and erosion processes in Redwood National Park in Humboldt County, California, with details provided in Data Releases by Iwatsubo *et al.* (18) and by Iwatsubo *et al.* (19). The study involved data collection at 53 sampling stations in the Redwood Creek and Mill Creek drainage basins. Although the stations were located in-stream, water quality probably represents runoff since many of the streams have very high slopes (typically up to about 25-30 percent) and have rocky, stable bottoms.

Water quality and hydrologic data collected include:

- Stream stage
- Stream discharge
- Turbidity
- Sediment characteristics
- Temperature*
- pH*
- Total alkalinity*
- Specific conductance*
- Dissolved oxygen*
- Major dissolved solids components
- Selected trace elements
- Nitrogen
- Phosphorus
- Organic carbon
- Fish
- Periphyton
- Phytoplankton
- Seston

---

*Measured in the field.

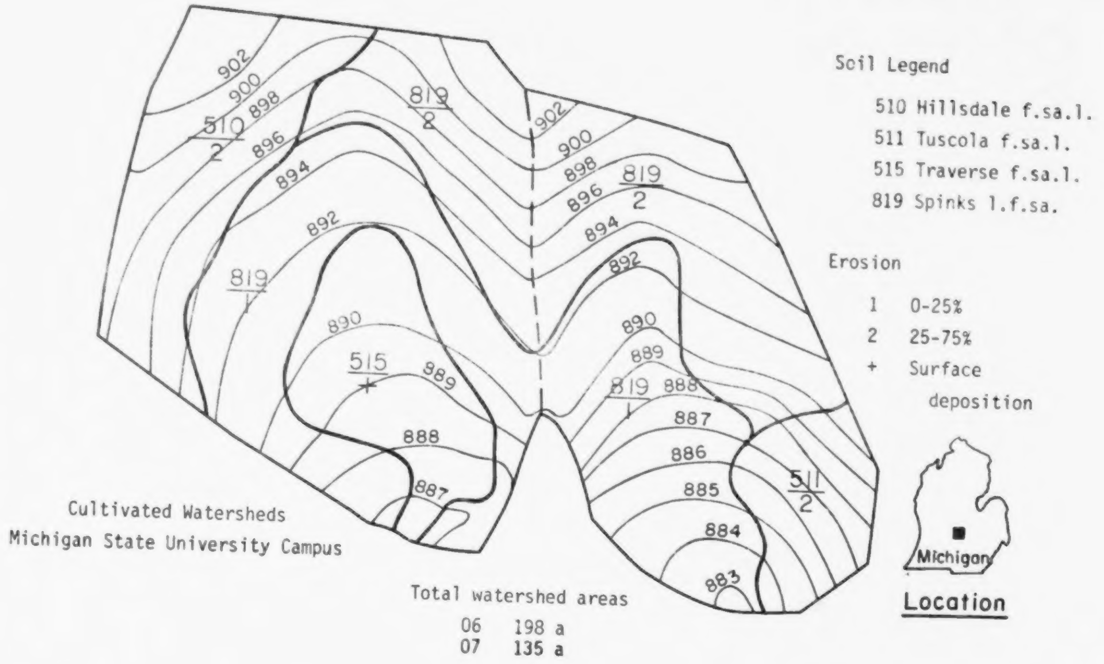

902
900
510
2
898
896
894
819
2
892
819
1
890
515
+
889
888
887
902
900
898
896
819
2
894
892
890
889
819
1
888
887
886
885
884
883
511
2
Cultivated Watersheds
Michigan State University Campus
Total watershed areas
06 198 a
07 135 a
Soil Legend
510 Hillsdale f.sa.l.
511 Tuscola f.sa.l.
515 Traverse f.sa.l.
819 Spinks l.f.sa.
Erosion
1 0-25%
2 25-75%
+ Surface deposition
Michigan
Location

Figure 12. Location and configuration of Michigan State University test plots. (Redrawn from Ellis (12))

- Organic carbon*
- Bacteria
- Pesticides*
- Benthic Invertebrates*

The study also examined changes in geometry at 10 channel cross sections on Mill Creek and the erosional landform distribution in that basin, quantity and chemical quality of rainwater, and streambed characteristics at certain Redwood Creek drainage basin stations.

For purposes of the present study, although the Redwood National Park data base is extensive, problems of analysis arose because suspended sediment analyses were not available for the samples for which chemical water quality data were developed. Although this is no criticism of the USGS studies and the situation probably arose because of the use of specialized sampling equipment, it did severely restrict the number of observations where the chemical water quality measurements and suspended solids concentrations could be considered to represent a single observation. In general, if the times of sampling were sufficiently close (within about 10 minutes), conditions were considered constant, and the data for suspended solids were combined with water quality data. Because of these limitations, however, there were few acceptable observations for analysis (see Table 8).

Seattle Data Base--

Runoff data for the City of Seattle, Washington, were obtained from a study performed for the Municipal Environmental Research Laboratory, U.S. Environmental Protection Agency. A brief description of the data base is provided in a document prepared by Huber and Heaney (20). Runoff water quality parameters available for analysis were:

- Orthophosphate P
- Total phosphorus
- Conductivity
- Ammonia N
- Suspended solids
- Nitrate plus nitrite N
- Organic nitrogen
- Lead
- Zinc
- Turbidity

The data base includes data for six different sampling sites in or near Seattle. In the present study, data from three sampling sites were subjected to analysis. These were Viewridge 1, South Seattle, and Southcenter. Figure 13 shows the location of the three sampling sites.

Honey Creek Data Base--

Runoff data for Honey Creek at Melmore, Ohio, were extracted from a U.S. Army Corps of Engineers report (21), which describes the sampling site as follows:

---

*Benthic analyses.

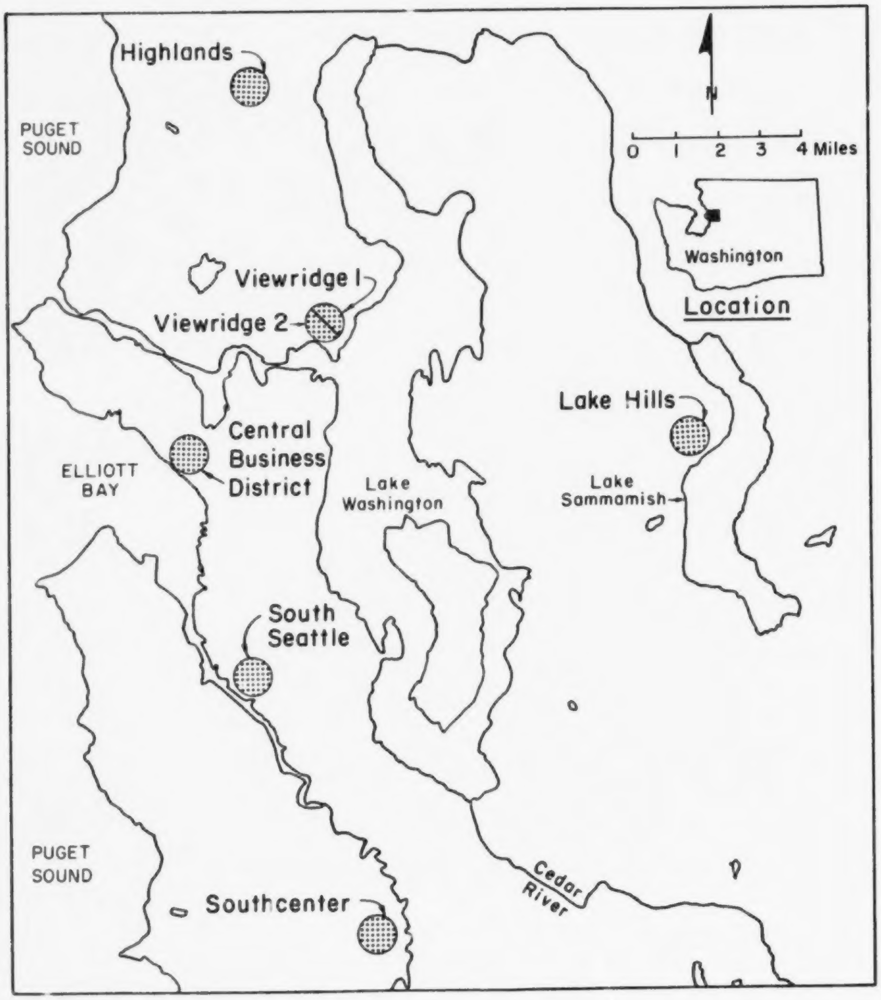

Highlands
PUGET SOUND
0 1 2 3 4 Miles
N
Washington
Location
Viewridge 1
Viewridge 2
Lake Hills
Central Business District
ELLIOTT BAY
Lake Washington
Lake Sammamish
South Seattle
PUGET SOUND
Southcenter
Cedar River

Figure 13. Location map for Seattle catchments.
(Redrawn from Huber and Heaney (20))

"04197100 HONEY CREEK AT MELMORE, OH

Lat 41°01'20", long 83°06'35", Seneca County, Hydrologic Unit 04100011, at bridge on State Highways 67 and 100 at Melmore, 1.5 mi (2.4 km) upstream from Buckeye Creek."

Available water quality parameters were:

- Total phosphorus P
- Orthophosphate P
- Nitrate plus nitrite N
- Ammonia N
- Total kjeldahl nitrogen
- Silicate
- Suspended solids
- Chloride
- Conductivity
- Iron

SECTION 5

RESULTS AND DISCUSSION

Results of individual analyses will be presented for each data base. First, correlations between suspended sediment and each water quality constituent will be discussed, followed by results of multiple regression analysis. In the discussion of simple regression results, tables of statistics are provided, with results being presented separately for each data base. A summary table is also presented showing estimated potency factors by watershed. In the discussion of multiple regression analyses results are organized around location, and specific water quality relationships with variables other than suspended solids are emphasized.

SIMPLE LINEAR REGRESSION

In all tables regarding simple linear regression, the notation described below is used. Assuming the equation (same form as equation 3):

$$C = b + mS$$

columns in each table include the slope of the line (m), the intercept (b), and the 95% confidence limits for the slope and intercept. The limits are denoted as m_ℓ (lower limit for the slope), m_u (upper limit for the slope) and corresponding b_ℓ and b_u. Note that where "m*" is to be used as a potency factor in the NPS model, it must be multiplied by 100%. See note on Table 23. The value for r^2, the F statistic, and the number of observations are also provided. The columns headed S_m, S_b, and S_F contain either blank or "X" entries. Where "X" is shown, the particular statistic (e.g., F for S_F) is significant at the $\alpha = 0.05$ level. That is, the slope, intercept, and/or correlation (S_m, S_b, and/or S_F) are non-zero. The tables also show approximate ranges for the dependent and independent variables as X_u, X_ℓ, Y_u, Y_ℓ (for the upper and lower limits of X and Y, respectively), and the number of observations, n.

Watkinsville

As discussed earlier, three separate data bases from Watkinsville, Georgia, were subjected to analysis. Results of simple linear regression for the relationship between suspended solids and various other runoff water quality constituents are shown in Tables 10 through 12.

TABLE 10. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATIONS (IN g/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF DISSOLVED SUBSTANCES IN RUNOFF) FROM WATKINSVILLE, GEORGIA. DATA ARE FROM PLOT P-04

Constituent, Units	m_u	m*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
**Atrazine, ppb	.744	.105	-.535		1.82	1.02	.227	X	.006	.117		2.17	.01	2.57	.27	21
TKN, ppm	1.73	1.30	.878	X	1.57	1.23	.886	X	.371	37.8	X	2.17	.01	4.31	.74	66
NH_4-N, ppm	1.19	.850	.506	X	.186	-.085	-.356		.185	24.0	X	2.17	.01	3.00	0.0	108
**NO_3-N, ppm	.814	.541	.268	X	-.601	-.821	-1.04	X	.197	15.7	X	2.17	.01	0.82	-1.02	66
PO_4-P, ppm	.119	.062	.005	X	.164	.118	.073	X	.069	4.78	X	2.17	.01	.500	.005	66
**Cl, ppm	.443	.247	.051	X	.323	.168	.014	X	.056	6.25	X	2.17	.01	1.59	-1.04	108

*Note that m is the potency factor for b not significantly different from zero at $\alpha = .05$ (S_b blank) and the dependent variable not transformed. If b is significantly non-zero, then it should be recognized that $C = mS$ was not a very acceptable linear relationship for the particular constituent given the available data. Where S_m is blank (m not significant) the potency factor m is very tenuous. The value of r^2 represents the proportion of the variance of the data for each constituent which may be accounted for by suspended sediment. For use with the NPS Model, m must be multiplied by 100%. Because of g/ℓ units for suspended solids at Watkinsville and Michigan, m for these sites should only be divided by 10.

**Log_{10} transformation. That is,

$$\log_{10} C = b + mS$$

Note that a very small incremental constant (on the order of the smallest observed value) was added to the data where logs were used to prevent log 0.

See text for interpretation of column headings.

TABLE 11. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATIONS (IN g/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF DISSOLVED SUBSTANCES IN RUNOFF). DATA ARE FROM WATKINSVILLE, GEORGIA TEST PLOT P-03

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Trifluralin, ppb	.912	.694	.475	X	4.30	3.04	1.79	X	.191	39.7	X	17.7	0.0	26.30	0.0	170
**Trifluralin, ppb	.062	.050	.038	X	.580	.511	.442	X	.290	68.5	X	17.7	0.0	1.42	0.0	170
Diphenamid, ppb	82.1	62.7	43.3	X	42.2	-32.7	-108.		.155	40.6	X	17.7	0.0	3384.	0.0	223
†Diphenamid, ppb	2.70	2.30	1.91	X	.391	.118	-.154		.378	134.	X	1.264	0.0	3.515	0.0	223

Note: See comments at foot of Table 10 and text for interpretation of column headings.

†Signifies both dependent and independent variables had been $\log_{10}$ transformed.

TABLE 12. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATIONS (IN g/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF DISSOLVED SUBSTANCES IN RUNOFF) DATA ARE FROM WATKINSVILLE, GEORGIA TEST PLOT P-01.

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
**Diphenamid, ppb	.0020	.0016	.0012	x	.772	.579	.387	x	.499	63.7	x	2046.	0.0	3.0	0.0	66
**Trifluralin, ppb	.031	.023	.016	x	.392	.277	.161	x	.505	43.9	x	58.4	.08	1.45	0.0	45

Note: See comments at foot of Table 10 and text for interpretation of column headings.

does not affect total water column loads of pollutants as they relate to suspended solids. Table 13 shows correlations between dissolved pollutants and pollutants in the suspended solids fraction for Watkinsville data.

Atrazine was not found to have a statistically significant relationship to suspended sediment concentration at $\alpha = .05$ (see Table 10). The slope was not statistically different from zero, and r^2 was low, so F was similarly low (below critical F for the degrees of freedom). The amount of data available from this set was small for atrazine, and this may account for not detecting a relationship. The fact that the intercept was significantly non-zero suggests that some of the atrazine may move in the dissolved state so that the water concentration is non-zero when suspended sediment is not present.

From Table 13, it is evident that despite the lack of a significant relationship between suspended solids and dissolved atrazine concentration (see Table 10), there is a fairly consistent partition coefficient relating dissolved and suspended atrazine. These outcomes may seem to conflict, initially, but they may be easily reconciled. The significant partition coefficient simply states that the water column atrazine concentration is related to the concentration of atrazine in the suspended fraction, but not significantly to the concentration of suspended matter itself. The relationship between suspended fraction atrazine concentration (i.e., dissolved atrazine) and suspended solids load also depends upon the relative availability of both atrazine and solids to wash off. If the ratio of mass of atrazine washing off to mass of solids washing off were constant, then the relationship between dissolved atrazine and suspended solids would undoubtedly be better. Other results of the analysis for atrazine showed that atrazine runoff was very dependent on time from cropping ($r^2 = 0.96$, $F = 453$). Solids runoff data, although not regressed on time, showed no such clearcut relationship over time upon visual inspection.

All of the remaining regressions of dissolved species concentrations on suspended solids concentrations were significant. This means that the trends apparent in the data were not the result of chance alone and that "m," appropriately corrected to account for transformations used, and taking non-zero intercepts into consideration, may be used as estimates of potency factors. Where the data were transformed, it may be desirable to estimate a series of potency factors by a piecewise-linear approximation to the curve, and to select those which are appropriate. Similarly, where "b" is significant, it might be desirable to add a constant to estimates in NPS or SWMM in the calibration and verification. This can easily be done directly in the computer code.

Regression results for TKN, NH_4, NO_3, PO_4, and Cl are as would be expected, with TKN giving the highest r^2. The highly soluble species (NH_4, NO_3, PO_4, and Cl) may only be related to suspended solids to the extent that both are washed off simultaneously. Examining the relationships in Table 13 suggests very strongly that this is so. There are no significant relationships (essentially inverse partition coefficients) for TKN, NH_4, or PO_4. The significant partition coefficients for atrazine, diphenamid, and trifluralin indicate that these are probably associated with particulate matter and that sorption phenomena are probably involved. The idea of simulating these

TABLE 13. REGRESSION OF DISSOLVED SPECIES ON CONCENTRATION IN SUSPENDED SOLIDS USING DATA FROM WATKINSVILLE, GEORGIA. FORM OF RELATIONSHIP: $C_D = b + mC_S$ WHERE C_D = DISSOLVED SPECIES CONCENTRATION, C_S = CONCENTRATION OF SPECIES IN SUSPENDED FRACTION. UNITS ARE CONSISTENT BETWEEN C_D AND C_S

Constituent, Units	m_u	m	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Atrazine, ppb†	3.10	2.44	1.77	X	-3.07	-4.67	-6.27	X	.754	58.3	X	3.00	1.84	2.57	.30	21
Atrazine, ppb	.425	.369	.313	X	-5.95	-20.8	-35.6	X	.815	177.	X	1000.	70.0	370.	2.0	42
TKN, ppm††	.157	-.350	-.857		4.91	3.24	1.56	X	.029	1.91		4.5	2.1	4.3	.74	66
TKN, ppm	5×10^{-6}	-3×10^{-5}	-7×10^{-5}		2.48	2.21	1.93	X	.044	2.93		35365.	120.	4.3	.74	66
Ammonia N, ppm	6×10^{-6}	-6×10^{-5}	-2×10^{-5}		.665	.508	.350	X	.010	1.06		96800.	0.0	3.0	0.0	108
PO_4-P, ppm&	7×10^{-6}	-2×10^{-5}	-4×10^{-5}		.212	.176	.140	X	.028	1.87		8587.	.33	.50	.005	66
PO_4-P, ppm&,††	.017	-.065	-.147		.587	.348	.108	X	.038	2.50		3.93	1.55	.50	.005	66
Diphenamid, ppb	.346	.229	.113	X	128.	61.4	-5.28		.064	15.1	X	2726.	0.0	3384.	0.0	223
Diphenamid, ppb**	9.6×10^{-4}	7.8×10^{-4}	6.1×10^{-4}	X	1.20	1.03	.866	X	.263	79.0	X	2726.	0.0	3.5	0.0	223
Trifluralin, ppb	.034	.027	.020	X	4.43	3.41	2.39	X	.266	60.9	X	780.	0.0	26.3	0.0	170
Trifluralin, ppb†	.374	.315	.256	X	.301	.192	.082	X	.400	112.	X	2.89	0.0	1.4	0.0	170
Diphenamid, ppb	.308	.192	.075	X	75.7	14.6	-46.5		.076	10.7	X	2024.	0.0	2867.	0.0	131
Diphenamid, ppb**	.002	.0016	.0012	X	.772	.579	.387	X	.499	63.7	X	2024.	0.0	3.00	0.0	66
Trifluralin, ppb**	.0026	.0004	-.002		.681	.543	.406	X	.004	.173		257.	0.0	1.45	0.0	45

** $\log_{10} C_D = b + m\,C_s$

† $\log_{10} C_D = b + m \log_{10} C_s$

†† $C_D = b + m \log_{10} C_s$

& $C_D = PO_4$, C_s = Total P

Note: See comments at foot of Table 10 and see text for explanation of column headings. Note that m in this table does not represent potency factors. Note also that the first seven lines in the table pertain to plot P-04, lines 8-11 pertain to plot P-03, and lines 12-14 pertain to plot P-01.

species as a function of suspended sediment is not entirely unrealistic. That the r^2 values are not nearly unity suggests that the nature of the sorption equilibrium is variable, probably a result of changes in the nature of the sediment and physical and chemical conditions.

Regressions of dissolved diphenamid on suspended solids were not only significant, but also gave similar results for the partition coefficient for both plots P-01 and P-03. Note in Table 13 that for untransformed variables ("Diphenamid, ppb") the slope and intercept ranges overlap for the two plots and the r^2 values are similar. The r^2 values were not rigorously tested to show a statistical difference, however. Where the diphenamid concentration was transformed ("diphenamid, ppb**"), confidence limits on the slope and intercept do not overlap but are close. Here, the r^2 values are different and are both substantially higher than for the untransformed diphenamid concentration.

The regressions of the log of dissolved trifluralin ("trifluralin, ppb**") on suspended solids gave r^2 values of 0.290 (plot P-03) and 0.505 (plot P-01). Note that for trifluralin, as well as diphenamid and other constituents shown in Table 10 (except NH_4) the intercept is significantly nonzero. Since the relationship in NPS and SWMM between suspended solids and other pollutants does not include a constant, this cannot be accounted for without changing the code. However, the desirability of adding a constant as opposed to simply assuming a zero intercept, and potency factor m, is not clear.

Buffalo Bill Watershed

Table 14 shows analytical results obtained with the Buffalo Bill Watershed data. The tabulated information is for regressions of various runoff water quality constituent concentrations on suspended solids concentrations.

In this data base, there were several different sampling locations. Because the main objective of this study is to estimate potency factors, because the data appeared fairly homogeneous spatially, and because there were limited numbers of observations at each sampling site, the data were pooled spatially. The net effect of this, however, should be to reduce the significance of relationships since it introduces another unaccountable source of variance into the data. Thus true correlations are likely to be at least as good as suggested by the results in Table 14.

As shown in Table 14, DDE, DDT (both examined in case other pesticides prove to have similar behavior), and nitrate were not correlated with suspended solids. The best correlations were with dieldrin, BOD, and fecal coliforms, although r^2 values for these were not high. It is instructive to compare results for the Buffalo Bill and Watkinsville watersheds. In general, results for constituents in common for the two data bases are very dissimilar. It should be noted, however, that those for which data were available in both data bases are also those for which correlation of concentrations between dissolved and suspended solids phases (Table 13) were weak. These are TKN, NH_4, NO_3, and PO_4. Clearly, though, even were data available for diphenamid,

TABLE 14. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATION (IN mg/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS IN RUNOFF) FROM THE BUFFALO BILL WATERSHED (SCOTT COUNTY, IOWA)

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Dieldrin, ppt[a]	.047	.040	.033	X	9.37	4.59	-.200		.597	123.	X	8563.	13.0	130.0	2.7	85
Dieldrin, ppt**	1.5×10^{-4}	1.3×10^{-4}	9.6×10^{-5}	X	1.05	.968	.883	X	.469	73.2	X	8563.	13.0	2.11	.43	85
DDE, ppt	5.3×10^{-4}	2.1×10^{-4}	-1.1×10^{-4}		7.05	6.10	5.15	X	.020	1.69		8563.	13.0	24.8	1.8	83
DDT, ppt	6.9×10^{-4}	3.2×10^{-4}	-5.1×10^{-5}		5.72	4.33	2.94	X	.071	3.04		8563.	13.0	17.9	2.7	42
DDT, ppt**	3.7×10^{-5}	1.4×10^{-5}	8.5×10^{-5}		.712	.627	.541	X	.038	1.59		8563.	13.0	1.25	0.44	42
Phosphate, ppm	-2.7×10^{-6}	-2.1×10^{-5}	-4.0×10^{-5}	X	.314	.261	.208	X	.056	5.21	X	8560.	5.6	1.09	.010	89
Phosphate, ppm**	-1.5×10^{-5}	-4.9×10^{-5}	-8.4×10^{-5}	X	-.608	-.707	-.805	X	.084	7.98		8560.	5.6	.04	-2.0	89
Nitrate-N, ppm	7.1×10^{-5}	-4.7×10^{-6}	-8.1×10^{-5}		1.52	1.31	1.09	X	1.7×10^{-4}	.015		8560.	5.6	4.07	.17	89
Nitrite-N, ppm	2.4×10^{-5}	1.5×10^{-5}	7.2×10^{-6}	X	.079	.055	.032	X	.138	13.9	X	8560.	5.6	.466	.015	89
Nitrite-N, ppm**	7.9×10^{-5}	4.9×10^{-5}	1.9×10^{-5}	X	-1.23	-1.32	-1.40	X	.105	10.2	X	8560.	5.6	-.332	-1.77	89
Ammonia N, ppm	9.6×10^{-5}	6.3×10^{-5}	3.1×10^{-5}	X	.421	.329	.237	X	.150	15.3	X	8560.	5.6	1.59	.037	89
TKN, ppm	5.8×10^{-4}	3.6×10^{-4}	1.4×10^{-4}	X	3.40	2.77	2.15	X	.111	10.8	X	8560.	5.6	9.47	.27	89
Fecal Colis, MPN/ 100 ml**	2.3×10^{-4}	1.7×10^{-4}	9.8×10^{-5}	X	4.34	4.15	3.95	X	.214	23.7	X	8560.	5.6	5.96	2.44	89
BOD, ppm	2.2×10^{-3}	1.7×10^{-3}	1.2×10^{-3}	X	8.67	7.18	5.69	X	.320	40.9	X	8560.	5.6	29.5	.94	89

[a] ppt = parts per trillion

** $\text{Log}_{10}\ C = b + mS$

Note: See comments at foot of Table 10 and text for explanation of column headings.

atrazine, and trifluralin on the Buffalo Bill Watershed, results might not compare well because of the availability factor discussed earlier. Thus further emphasizes the highly site-specific nature of the potency factor.

Michigan State Farms

In a study paralleling that at Watkinsville, Georgia, runoff data were collected from test plots near East Lansing, Michigan. Table 15 shows results of regression analyses using the data from test plot P-06 from the Michigan State University study.

Regression results with the Michigan State University data showed very weak relationships between suspended solids and all other runoff quality constituents, although statistically significant relationships were obtained for all except nitrate. The best relationships were with atrazine in a log-log regression ($r^2 = 0.138$) and total phosphorus ($r^2 = 0.187$). With respect to TKN and ammonia, relationships are significant but borderline.

Comparing these results with those from the Watkinsville data (Table 10), the atrazine confidence limits for slopes and intercepts are similar. In contrast, however, those for TKN and ammonia have none of their ranges in common.

Some comments should be made here regarding the Michigan State data base. First, regarding comparison with the Watkinsville data, climatic factors should be considered. During part of the year, precipitation at the Michigan site is in the form of snow, and the ground is frozen. Secondly, several large-scale errors were noted in the data base such as mispunched values and off-column data. The fact that these errors were found leads to some suspicion about the veracity of the remaining data, and although every effort was made to purge errors from the data base in the present study, results presented in Table 15 should be used with appropriate caution.

Redwood National Park

As discussed earlier, the data base for Redwood National Park is very large, but data for suspended solids were not obtained for the same samples as were water quality data. Accordingly, observations from which correlations could be computed were inadequate. An attempt was made to pool data both spatially and temporally, justifying this by the relatively uniform geology, land-use, and soil type and requiring suspended sediment to have been observed within 10 minutes of other water quality parameters. Nonetheless, it was possible to obtain only from 7 to 17 observations, and regressions were either non-significant or borderline-significant. In view of the data limitations, non-significant fitted lines, and likely misleading results of analyses, the statistics will not be presented.

Seattle: Viewridge 1

Data for three distinct urban watersheds in Seattle, Washington, were subjected to regression analysis. The first to be discussed here is Viewridge 1. As described in Table 8, this watershed represents a high density urbanized area containing many old residences. Table 16 shows results of regressions of runoff water quality on runoff suspended solids concentration.

TABLE 15. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATION (IN g/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF SUBSTANCES IN RUNOFF) FROM MICHIGAN STATE UNIVERSITY STUDY

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Atrazine, ppb**	.107	.060	.013	X	1.52	1.44	1.35	X	.049	6.36	X	8.06	.084	2.43	0.0	125
Atrazine, ppb†	.431	.298	.165	X	1.62	1.56	1.49	X	.138	19.8	X	.895	-1.06	2.43	0.0	125
TKN, ppm	.653	.271	-.112		5.33	4.53	3.74	X	.021	1.58		8.07	0.18	11.02	0.90	74
TKN, ppm†	.321	.187	.054	X	.674	.623	.572	X	.080	6.23	X	.910	-.74	1.04	-.046	74
Ammonia-N, ppm	.299	.164	.029	X	2.08	1.80	1.53	X	.058	4.65	X	8.07	0.18	4.49	0.31	78
Ammonia-N, ppm**	.082	.048	.014	X	.253	.183	.114	X	.075	6.19	X	8.07	0.18	.652	-.510	78
Ammonia-N, ppm†	.316	.193	.070	X	.299	.253	.206	X	.092	7.74	X	.895	-.686	.652	-.510	78
Total P, ppm	.076	.053	.030	X	.476	.428	.380	X	.187	16.6	X	8.07	0.18	.798	.093	74
Total P, ppm**	.077	.051	.026	X	-.351	-.404	-.458	X	.148	12.5	X	8.07	0.18	-.113	-1.02	74
Total P, ppm†	.284	.191	.099	X	-.294	-.330	-.366	X	.157	13.4	X	.895	-.686	-.113	-1.02	74
Nitrate-N, ppm	.157	-.020	-.198		1.92	1.55	1.18	X	.0006	.043		8.07	0.18	5.12	0.17	74
Nitrate-N, ppm**	.081	.028	-.025		.117	.006	-.104		.012	.873		8.07	0.18	.697	-.729	74
Nitrate, ppm†	.346	.154	-.037		.120	.046	-.028		.028	2.04		.895	-.686	.697	-.729	74

**$\log_{10} C = b + mS$

†$\log_{10} C = b + m \log_{10} S$

Note: See comments at foot of Table 10 and text for explanation of column headings.

TABLE 16. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATION (IN mg/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF DISSOLVED AND SUSPENDED SUBSTANCES IN RUNOFF) FROM VIEWRIDGE 1, SEATTLE, DATA

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Ortho PO_4-P, ppm†	.255	.172	.089	X	-1.31	-1.43	-1.55	X	.083	16.9	X	3.66	0.0	.012	-2.01	188
Total P, ppm†	.311	.237	.162	X	-.807	-.917	-1.03	X	.175	39.5	X	3.66	0.0	.616	-1.23	188
Ammonia N, ppm†	.481	.336	.191	X	-1.83	-2.05	-2.26	X	.102	21.0	X	3.66	0.0	.108	-2.0	188
NO_2 + NO_3-N, ppm†	-.047	-.140	-.233	X	.030	-.108	-.246		.046	8.92	X	3.66	0.0	.63	-1.49	188
Org N, ppm†	.348	.280	.212	X	-.346	-.447	-.547	X	.263	66.5	X	3.66	0.0	.798	-2.04	188
Conductivity, μmhos/cm††	-11.1	-24.6	-38.2	X	164.	144.	124.	X	.065	13.0	X	3.66	0.0	246.8	26.0	188
Turbidity, JTU†	.432	.358	.284	X	.646	.537	.428	X	.332	92.5	X	3.66	0.0	2.00	-.968	188
Lead, ppb†	.572	.490	.408	X	1.54	1.42	1.30	X	.430	140.1	X	3.66	0.0	3.22	1.28	188
Zinc, ppb†	.347	.283	.218	X	1.57	1.47	1.38	X	.289	75.5	X	3.66	0.0	2.77	.979	188

†$\log_{10} C = b + m \log_{10} S$

††$C = b + m \log_{10} S$

Note: See comments at foot of Table 10 and text for explanation of column headings.

All runoff water quality constituents were significantly correlated with suspended solids. In some cases, notably lead, zinc, turbidity, and organic nitrogen, values of r^2 were substantial with r^2 being as high as 0.43 for lead. Thus 43% of the variance of runoff lead concentration can be explained by the suspended solids concentration. Again, it must be emphasized that this may be in part due to association of lead with particulate matter, but it may also be due at least in part, to some concomitant factor. In reality, the strong correlation probably represents both physical association and concomitant variables causing both lead and suspended matter to be flushed from the watershed.

Zinc was also found to be strongly correlated with suspended solids ($r^2 = 0.289$) and as well, with lead in the runoff ($r^2 = 0.557$ for $\log_{10} [Z_n] = b + m \log [P_b]$). As might be expected, total phosphorus was correlated more strongly with suspended solids than was orthophosphate, which is highly soluble, and would be expected to be less associated with suspended solids. Even the sum of nitrate and nitrite (as a single parameter), which is very soluble and generally considered to be unassociated with particulate matter, was significantly correlated with suspended solids. Perhaps it is reasonable to assume that this correlation ($r^2 = 0.046$) represents that purely due to concomitant factors such as the propensity of overland flow to carry both pollutants and suspended solids when they are not associated.

It is of interest to compare correlations shown in Table 16 with regressions on turbidity. These are shown in Table 17. Despite the fact that the correlation between turbidity and suspended solids is not very strong ($r^2 = 0.332$), the values of r^2 in Table 17 correspond very well in all cases with those in Table 16. Apparently, the variance shared in common between turbidity and suspended solids is also essentially variance shared with other pollutants.

It is also of interest to examine the correlations between runoff water quality parameters and the number of antecedent dry days. This is the case because some of the runoff models (e.g., SWMM) predicate quality on "dust and dirt" accumulation, which is, in turn, a function of antecedent dry days. Table 18 shows these correlations.

The correlation between suspended solids concentrations and number of antecedent dry days is significant ($r^2 = 0.114$, $F = 23.9$). However, with only three exceptions (conductivity, $NO_2 + NO_3$, and organic nitrogen), direct correlations between pollutants in the runoff and number of antecedent dry days are higher than between suspended solids and antecedent dry days for this data base. This suggests that it might be better to estimate many pollutants in urban runoff directly as a function of dry days without introducing the suspended solids intermediate. This is especially true for orthophosphate, total phosphorus, and ammonia which, while correlating well with number of antecedent dry days, correlated very poorly with suspended solids. Since they are poorly correlated with suspended solids, and because suspended solids is, in turn, poorly correlated with number of antecedent dry days, this means the algorithm as used in SWMM can be especially poor for simulating these pollutants exported from urban sheds. It must be emphasized, however, that this is true for this data base and other watersheds do not suggest this as strongly.

TABLE 17. STATISTICS FOR CORRELATIONS BETWEEN TURBIDITY (IN JACKSON TURBIDITY UNITS) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF SUBSTANCES IN RUNOFF) FROM VIEWRIDGE 1, SEATTLE, DATA

Constituent, Units	m_u	m	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Ortho PO_4-P, ppm[†]	.420	.287	.154	X	-1.35	-1.49	-1.63	X	.089	18.2	X	2.04	-.968	.012	-2.01	188
Total P, ppm[†]	.531	.413	.295	X	-.890	-1.02	-1.15	X	.206	48.3	X	2.04	-.968	.616	-1.32	188
Ammonia N, ppm[†]	.700	.463	.226	X	-1.81	-2.07	-2.32	X	.075	15.0	X	2.04	-.968	.108	-2.01	188
NO_2 + NO_3-N, ppm[†]	.047	-.106	-.259		-.025	-.190	-.356	X	.010	1.90		2.04	-.968	.600	-1.70	188
Org-N, ppm[†]	.571	.463	.355	X	-.423	-.541	-.658	X	.279	71.9	X	2.04	-.968	.924	-2.04	188
Conductivity, μmhos/cm[††]	-.034	-22.4	-44.8	X	158.	134.	109.	X	.021	3.95	X	2.04	-.968	246.8	26.0	188
Lead, ppb[†]	.877	.739	.601	X	1.48	1.33	1.18	X	.377	113.	X	2.04	-.968	3.22	1.28	188
Zinc, ppb[†]	.527	.420	.313	X	1.54	1.43	1.31	X	.246	60.8	X	2.04	-.968	2.77	.979	188

[†] $\log_{10} C = b + m \log_{10} S$

[††] $C = b + m \log_{10} S$

Note: See comments at foot of Table 10 and see text for explanation of column headings.
Note that m in this table does not represent potency factors.

TABLE 18. STATISTICS FOR CORRELATIONS BETWEEN ANTECEDENT DRY DAYS (DD) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF DISSOLVED AND SUSPENDED SUBSTANCES IN RUNOFF) FROM VIEWRIDGE 1, SEATTLE, DATA

Constituent, Units	m_u	m	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Ortho PO_4-P, ppm[†]	.355	.288	.221	X	-1.27	-1.32	-1.37	X	.279	72.0	X	1.59	-1.02	.012	-2.01	188
Total P, ppm[†]	.359	.298	.236	X	-.675	-.721	-.767	X	.332	92.5	X	1.59	-1.02	.616	-1.32	188
Ammonia N, ppm[†]	.644	.523	.403	X	-1.73	-1.82	-1.91	X	.286	74.6	X	1.59	-1.02	.108	-2.0	188
NO_2 + NO_3-N, ppm[†]	.140	.053	-.034		-.256	-.321	-.387	X	.008	1.48		1.59	-1.02	.63	-1.49	188
Org-N, ppm[†]	.188	.118	.047	X	-.063	-.116	-.169	X	.056	11.0	X	1.59	-1.02	.798	-2.04	188
Conductivity, μmhos/cm[††]	8.59	-4.24	-17.1		122.	113.	103.	X	.002	.430		1.59	-1.02	246.8	26.0	188
Turbidity, JTU[†]	.278	.200	.123	X	.999	.940	.882	X	.125	26.5	X	1.59	-1.02	2.00	-.968	188
Lead, ppb[†]	.458	.376	.293	X	1.99	1.92	1.86	X	.302	80.6	X	1.59	-1.02	3.22	1.28	188
Zinc, ppb[†]	.237	.171	.106	X	1.83	1.78	1.73	X	.127	27.1	X	1.59	-1.02	2.77	9.79	188
Suspended solids, ppm[†]	.434	.309	.184	X	1.33	1.23	1.14	X	.114	23.9	X	1.59	-1.02	3.66	0.0	188

[†] $\log_{10} C = b + m \log_{10} DD$

[††] $C = b + m \log_{10} DD$

Note: See comments at foot of Table 10 and see text for explanation of column headings. Note that m in this table does not represent potency factors.

There are other implications of antecedent dry days correlations as well. It might be interpreted that those pollutants correlating well with antecedent dry days but not with suspended solids do, in fact, accumulate on urban watersheds over time, but are not strongly associated with particulate matter. In this case, suspended solids do not appear to accumulate on the watershed as a simple function of time, at least on the time scale represented in the Viewridge 1 data.

In contrast, some pollutants not correlating well with suspended solids also do not correlate well with number of antecedent dry days. This is true for $NO_3 + NO_2$. Apparently, as suggested by this data base, these pollutants not only do not associate with particulate matter, they also do not appear to accumulate as a function of dry days. On the other hand, lead and zinc correlate both with suspended solids and with dry days suggesting time-based accumulation and possible association with particulate. Turbidity, as would be expected, correlates well with suspended solids, and like suspended solids, correlates only weakly with number of antecedent dry days.

Seattle: South Seattle

The South Seattle Watershed is an industrialized urban area, geographically located to the south of Viewridge 1. Table 8 summarizes the data base for this watershed, and its location is shown in Figure 13. Table 19 shows results of statistical analysis of runoff data.

For several constituents (orthophosphate, ammonia, and conductivity; see note a in Table 19), autocorrelation in the residuals suggested that observations could not be considered independent. The statistics are shown only for comparison purposes, as taking first differences appears not to be satisfactory under conditions exhibited by this data base (see Appendix A and note a in Table 19).

For constituents not transformed to first differences (the five at the top of Table 19) some of the correlations with suspended solids were very strong. These are lead ($r^2 = 0.671$), total P ($r^2 = 0.505$), and zinc ($r^2 = 0.390$). Comparing with the corresponding entries in Table 16 (Viewridge 1), the squared correlation coefficients (r^2) are 0.430, 0.175, and 0.289, respectively--all lower, with total P substantially so. It should be pointed out that these comparisons of r^2 are subjective and do not represent a statistical test. Statistical comparisons of r^2 were not made in this study.

Similarly, for $NO_2 + NO_3$ the correlation appears stronger than for the Viewridge 1 data. For organic nitrogen, the relationship is reversed, with the correlation weaker than for Viewridge 1.

Turning now to the 95% confidence limits on the slope (m) and intercept (b), those for Total P and $NO_2 + NO_3$ are different comparing South Seattle and Viewridge 1 (Tables 19 and 16). This is a somewhat subjective comparison made by examining the ranges between m_u and m_ℓ and b_u and b_ℓ for overlap without making any statements about probabilities that the slopes and intercepts are in fact different.

TABLE 19. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS CONCENTRATION (IN mg/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS IN RUNOFF) FROM THE SOUTH SEATTLE DATA BASE

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Total P, ppm†	.478	.419	.360	X	-1.30	-1.40	-1.49	X	.505	201.	X	3.47	0.0	.409	-1.31	199
NO_2 + NO_3-N, ppm†	.349	.252	.156	X	-.806	-.967	-1.13	X	.121	27.1	X	3.47	0.0	.654	-1.55	199
Org. N, ppm†	.293	.211	.129	X	-.429	-.567	-.704	X	.116	25.8	X	3.47	0.0	.987	-2.04	199
Lead, ppb†	.546	.497	.448	X	1.50	1.42	1.33	X	.671	401.	X	3.47	0.0	3.12	1.27	199
Zinc, ppb†	.301	.256	.211	X	1.99	1.91	1.84	X	.390	126.	X	3.47	0.0	3.01	1.79	199
									see note a below							
Ortho PO_4-P, ppm†	.207	.132	.057	X	.043	.002	-.038		.063	12.3	X	1.95	-1.49	1.47	-.918	184
Ammonia N, ppm†	.202	.106	.010	X	.051	-.002	-.054		.026	4.80	X	1.95	-1.49	1.71	-1.14	184
Conductivity, µmhos/cm††	18.7	1.60	-15.5		3.63	-5.67	-15.0		1.8×10^{-4}	.034		1.95	-1.49	372.	-410.	184

† $\log_{10} C = b + m \log_{10} S$

†† $C = b + m \log_{10} S$

Note a: Regressions for constituents listed below this point were done on first differences rather than on the observations themselves. This was because of a problem of autocorrelated residuals. Results are presented here primarily for comparison purposes since they are not useful for developing potency factors. Also, as discussed briefly in Appendix A, taking first differences (i.e., regressing $Y_{i-1} - Y_i$ on $X_{i-1} - X_i$ rather than Y_i on X_i) can cause a severe and unrealistic reduction in r^2 under some circumstances. Please also see comments at foot of Table 10 and see text for explanation of column headings.

If the ranges do overlap, then they share values in common. Since both slopes and/or intercepts could have identical values, we cannot conclude that the slopes and/or intercepts are different for the fitted lines being compared. The confidence limits for both m and b for lead, comparing South Seattle and Viewridge 1, are almost exactly coincident. Considering that the ranges are fairly narrow and r^2 values are very high, this suggests strong and consistent relationships between lead and suspended solids for at least two urban watersheds.

Zinc exhibits similar confidence limits for m, comparing the two watersheds. The ranges for b appear to be different, however. Despite the relatively low r^2 for organic N at South Seattle, the confidence limits for both m and b overlap those at Viewridge 1, suggesting some spatially consistent washoff behavior of organic N in addition to washoff behavior for lead and zinc on these two watersheds.

Turning, now, to correlations with antecedent dry days, Table 20 shows correlations between this parameter and runoff quality. Correlations here are not nearly as strong as at Viewridge 1, possibly suggesting different, or at least more variable pollutant accumulation phenomena. In this data base, suspended sediment correlates more weakly with antecedent dry days than at Viewridge 1. Using the model:

$$\mathrm{Log}_{10}\ S = b + m \log_{10} DD \qquad (6)$$

$r^2 = 0.098$ (not shown in Table 20). For the same relationship with S representing turbidity instead of suspended solids, $r^2 = 0.045$. Thus, despite the fact that pollutant correlations with dry days are weaker here than at Viewridge 1, several of them (Total P, $NO_2 + NO_3$, and zinc) might still be better estimated directly from antecedent dry days than from suspended solids, even given the strong correlation between suspended solids and total P and zinc. It is interesting to note that at South Seattle, unlike at Viewridge 1, $NO_3 + NO_2$ was fairly strongly correlated with antecedent dry days. Organic N was not well correlated at either site.

Seattle: Southcenter

Southcenter is the third and final Seattle watershed to be examined. As shown in Table 8, it contains a new shopping center and is urbanized. See Figure 13 for the location of the sampling site. Table 21 provides statistics for correlations of suspended solids with various other pollutants.

As shown in Table 21, all regressions were statistically significant. That is, the trends have not occurred by chance alone. Figure 14 shows a comparison of the results from Southcenter, South Seattle, and Viewridge 1.

As shown in both Table 20 and Figure 14, the confidence limits for the slopes for Southcenter all overlap those for South Seattle except Organic N, while comparing slopes between Southcenter and Viewridge 1, the ranges for Total P, lead, and zinc overlap. In terms of intercepts, and comparing Southcenter and South Seattle, the confidence limits overlap for all but organic N

TABLE 20. STATISTICS FOR CORRELATIONS BETWEEN ANTECEDENT DRY DAYS (DD) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF SUBSTANCES IN RUNOFF) FROM SOUTH SEATTLE DATA BASE

Constituent, Units	m_u	m	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	$\bar{X}_u$	$\bar{X}_\ell$	$\bar{Y}_u$	$\bar{Y}_\ell$	n
Total P, ppm†	.265	.190	.115	X	-.779	-.841	-.904	X	.113	25.0	X	1.59	-1.02	.409	-1.56	199
NO_2 + NO_3-N, ppm†	.457	.375	.292	X	-.690	-.760	-.829	X	.209	80.3	X	1.59	-1.02	.654	-1.56	199
Org. N, ppm†	.234	.153	.072	X	-.248	-.316	-.383	X	.066	14.0	X	1.59	-1.02	.987	-1.53	199
Lead, ppb†	.257	.179	.100	X	.216	.210	.203	X	.094	20.5	X	1.59	-1.02	3.12	1.272	199
Zinc, ppb†	.189	.137	.085	X	2.29	2.24	2.20	X	.121	27.2	X	1.59	-1.02	3.03	1.79	199

†$\log_{10} C = b + m \log_{10} DD$

Note: See comments at foot of Table 10 and see text for explanation of column headings. Note that m in this table does not represent potency factors.

TABLE 21. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS (IN mg/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS IN RUNOFF) FROM SEATTLE SOUTHCENTER DATA BASE

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Total P, ppm†	.416	.312	.208	X	-1.09	-1.23	-1.37	X	.152	35.4	X	3.06	0.0	.644	-2.04	199
NO_2 + NO_3-N, ppm†	.443	.305	.166	X	-.098	-1.17	-1.36	X	.089	19.2	X	3.06	0.0	.987	-2.04	199
Org. N, ppm†	.620	.502	.383	X	-.867	-1.03	-1.19	X	.265	70.9	X	3.06	0.0	.987	-2.04	199
Lead, ppb†	.728	.629	.530	X	1.59	1.45	1.32	X	.446	159.	X	3.06	0.0	3.43	0.99	199
Zinc, ppb†	.422	.312	.203	X	1.80	1.65	1.50	X	.140	32.0	X	3.06	0.0	3.06	0.99	199

$^\dagger \log_{10} C = b + m \log_{10} S$

Note: See comments at foot of Table 10 and see text for explanation of column headings.

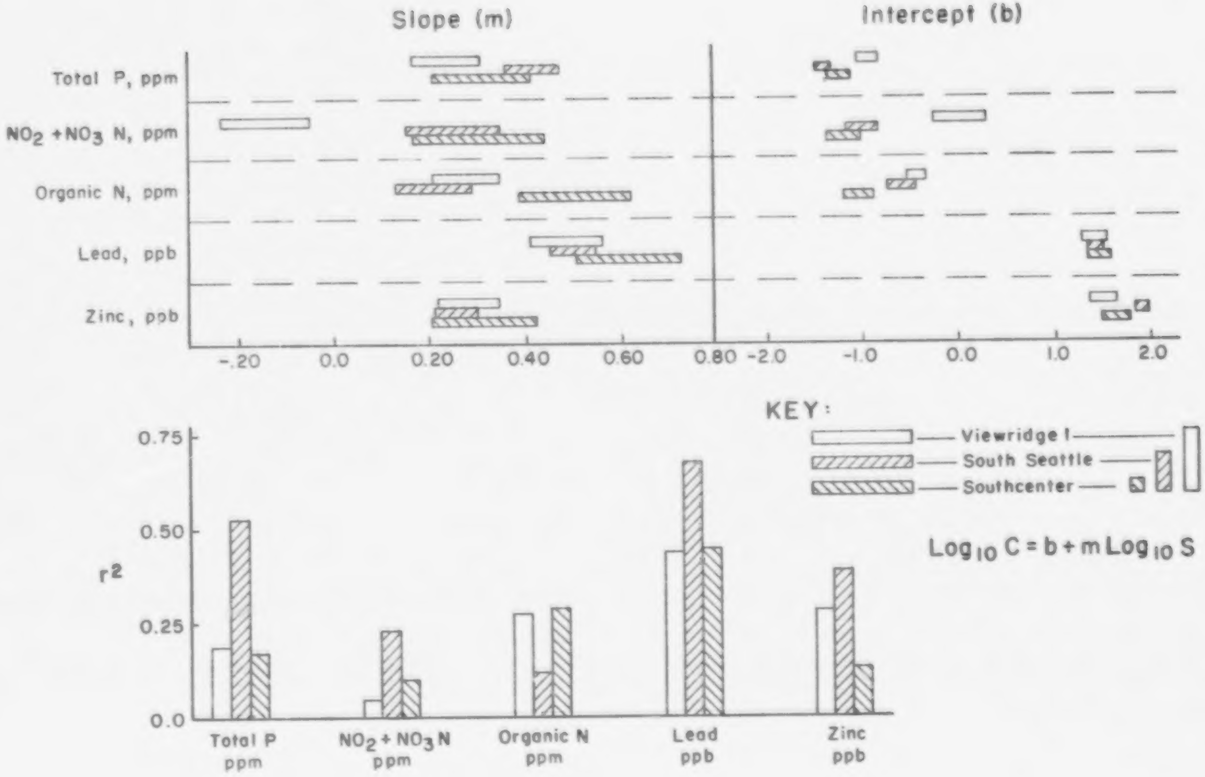


Figure 14. A comparison of statistical results among Southcenter, South Seattle, and Viewridge 1.

and zinc, although for zinc, they are close. Comparing confidence limits for intercepts between Southcenter and Viewridge 1, only those for lead and zinc overlap.

The ramifications of these comparisons are that if confidence limits are subjectively compared, and if it is assumed that where confidence limit estimates for two distinct sites share any values in common, they can be considered to be indistinct, then results for Seattle data are very consistent spatially. This is especially true for slopes, which are of particular interest here as potency factors. Although the data were not pooled and analyzed, a good estimate of the potency factor for any pollutant would be one shared in common by two or even all three sampling sites, bearing in mind, of course, that values from another urban area or even another watershed in Seattle might be somewhat different. The fact that the three sites in Seattle give such consistent results is, however, very encouraging. Later in this report, results of some multiple regression analyses are presented which further support (again in a subjective way) the contention that the three watersheds in Seattle behave very similarly in terms of runoff quality and its relationship to various other factors.

Honey Creek

The Honey Creek basin data base had the greatest number of observations of any examined in this study. It also provided satisfactory data for the greatest number of runoff water quality constituents, and because many storms were monitored with relatively few observations in each (rather than a few storms intensively sampled), there was also no residuals autocorrelation problem.

Table 22 shows correlations between suspended solids and other runoff water quality constituents for Honey Creek.

With the exception of orthophosphate, all regressions were significant. Many had very high r^2 values, some being the highest obtained in the study. These include iron, total phosphorus, and total phosphorus minus orthophosphate phosphorus. The fact that the F values are so high reflects both significance of the regressions and the large sample size.

A comment is in order here about the higher r^2 values observed in regression of the non-transformed over the log transformed data. Despite this, the transformed data regressions are probably more reliable since the resulting data distributions over both the X and Y axes are much more uniform while being similarly linear. In the untransformed data, much of the data formed a cluster with a few remote values strongly biasing r^2. Figures 15 and 16 are examples of plots showing this. Figure 15 is the Honey Creek regression of runoff iron concentration on suspended solids. Figure 16 shows the regression of $\log_{10}$ iron on $\log_{10}$ suspended solids. The data presented at the foot of the figures correspond to those presented in Table 22. Definitions for these statistics are provided in Section 3. The plotted numbers represent the number of observations at each set of coordinates. The line of "+" symbols is the fitted line, whereas the line of "•" symbols represents the confidence

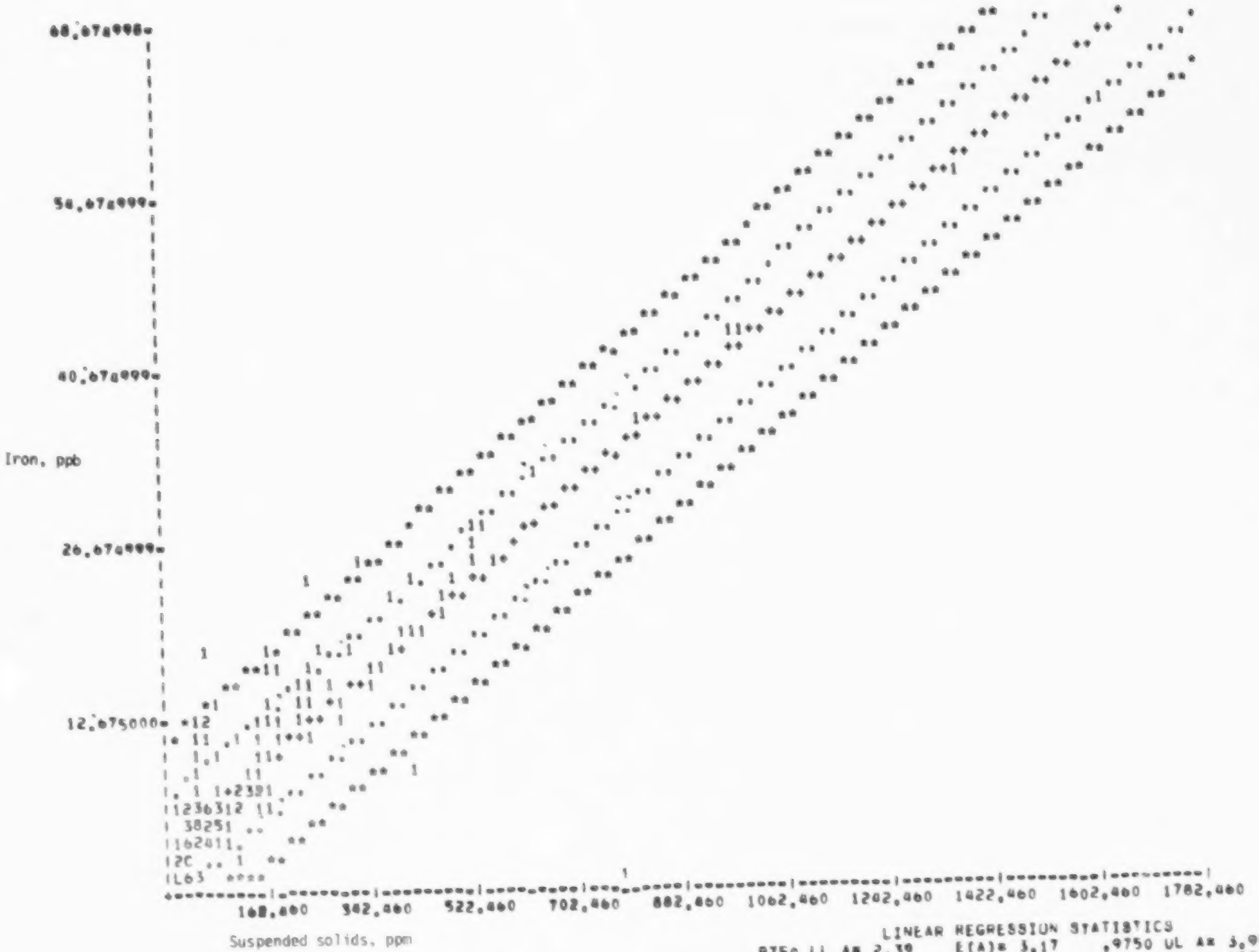
TABLE 22. STATISTICS FOR CORRELATIONS BETWEEN SUSPENDED SOLIDS (IN mg/ℓ) AND CONSTITUENTS SHOWN (CONCENTRATIONS OF SUBSTANCES IN RUNOFF) FROM THE HONEY CREEK WATERSHED

Constituent, Units	m_u	m^*	m_ℓ	S_m	b_u	b	b_ℓ	S_b	r^2	F	S_F	X_u	X_ℓ	Y_u	Y_ℓ	n
Total P, ppm†	.432	.405	.378	X	-1.26	-1.31	-1.36	X	.577	864.	X	3.21	-.740	0.156	-2.00	635
Total P, ppm	1.2×10^{-3}	1.1×10^{-3}	1.0×10^{-3}	X	.194	.183	.172	X	.679	1340.	X	1620.	.479	1.43	0.0	635
Ortho PO_4-P, ppm†	.062	.021	-.021		-1.08	-1.15	-1.23	X	.001	.929		3.21	-.740	0.154	-2.00	628
$NO_2 + NO_3$-N, ppm†	.256	.214	.171	X	.342	.266	.189	X	.134	98.1	X	3.21	-.740	1.24	-0.37	636
Ammonia N, ppm†	-.080	-.155	-.230	X	-.609	-.745	-.880	X	.026	16.4	X	3.21	-.740	.295	-2.24	608
TKN, ppm	2.9×10^{-3}	2.1×10^{-3}	1.4×10^{-3}	X	2.27	2.01	1.75	X	.194	29.0	X	1620.	.479	7.46	0.56	123
TKN, ppm†	.241	.182	.123	X	.093	-.027	-.146		.238	37.9	X	3.21	-.740	.864	-.240	123
Chloride, ppm†	-.148	-.172	-.195	X	1.81	1.77	1.73	X	.254	208.	X	3.21	-.740	2.11	.535	612
SiO_2, ppm†	.107	.076	.045	X	.734	.679	.624	X	.090	24.0	X	3.21	-.740	1.11	.049	244
Iron, ppb	.042	.040	.037	X	3.95	3.17	2.39	X	.828	813.	X	1620.	.479	56.1	.075	171
Iron, ppb†	.897	.815	.733	X	-.648	-.807	-.966	X	.697	388.	X	3.21	-.740	1.78	-1.14	171
Conductivity, µmhos/cm††	-.134	-.152	-.170	X	2.97	2.94	2.91	X	.310	284.	X	3.21	-.740	3.04	2.28	633
Total P-Ortho PO_4, ppm	1.2×10^{-2}	1.1×10^{-2}	1.0×10^{-2}	X	.096	.086	.077	X	.752	1897.	X	1620.	.479	1.78	0.0	627
Total P-Ortho PO_4, ppm†	.554	.474	.394	X	-1.49	-1.64	-1.78	X	.180	135.	X	3.21	-.740	2.91	-3.07	619

† $\log_{10} C = b + m \log_{10} S$

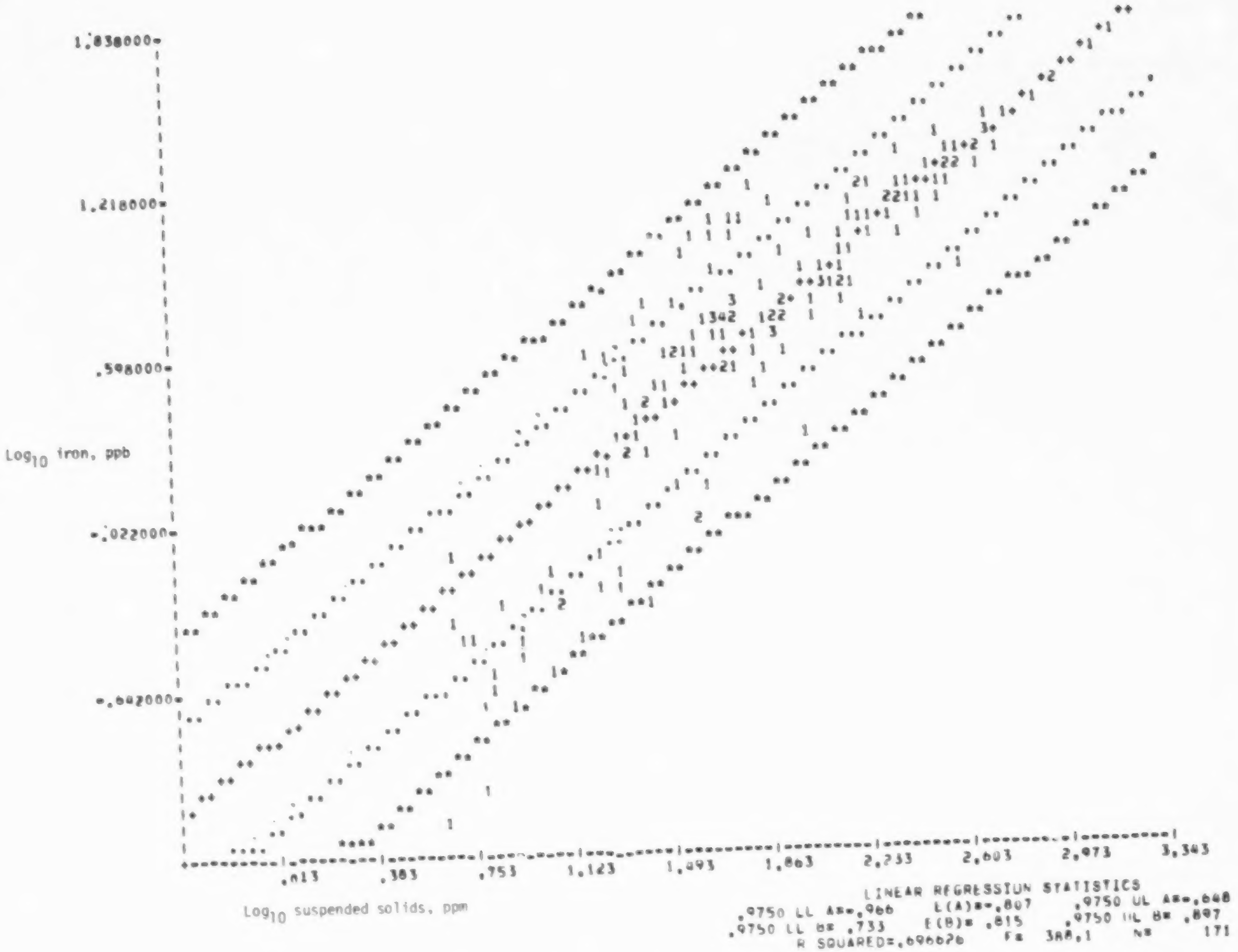
†† $C = b + \log_{10} S$

Note: See comments at foot of Table 10 and see text for explanation of column headings.



```
                  LINEAR REGRESSION STATISTICS
  ,9750 LL A= 2,39     E(A)= 3,17     ,9750 UL A= 3,95
,9750 LL B= ,368-01 E(B)= ,395-01 ,9750 UL B= ,423-01
       R SQUARED=,827930      F= 813,2     N=     171
```

Figure 15. Scatter plot and regression of iron concentration on suspended solids concentration.



```
                     LINEAR REGRESSION STATISTICS
 .9750 LL A=-.966    E(A)=-.807     .9750 UL A=-.648
.9750 LL B= .733    E(B)= .815     .9750 UL B= .897
        R SQUARED=.696626      F=  388.1     N=      171
```

Figure 16. Scatter plot of $\log_{10}$ iron concentration on $\log_{10}$ suspended solids concentration.

interval estimate of the population mean for a given value on the abscissa. The line of "*" symbols is the corresponding confidence interval estimate for prediction of individual values on the ordinate axis.

Table 23 presents a summary of potency factors derived from the results of this study.

MULTIPLE REGRESSION

As shown in the preceding section, suspended sediment concentration can generally account for only a small proportion of the variability of other runoff water quality constituents. Multiple regression analysis was used to examine other variables in concert as predictors of runoff quality. The variables included those shown in Table 24.

In the regression analyses, it was decided to include water quality constituents as independent variables since the purpose of multiple regression analysis here was less for prediction than for explication. Each step in multiple regression amounts to partialing out the independent variable from the dependent variable (see Appendix A) such that the array of residuals is a vector orthogonal to (independent of) each independent variable already used in the regression. That is, the coefficient of linear correlation between the residuals (difference between observed and predicted values of the dependent variable) and each of the predictor variables is identically zero. Examining multiple regression results with respect to variables statistically correlating with the residuals at each step allows consideration of possible physical phenomena and mechanisms influencing the values of the dependent variable.

Watkinsville: Plot P-04

Table 25 shows correlations (r) between dependent and independent variables for Watkinsville test plot P-04 (1974) data. Since not all variables were used as predictors for all regressions, some correlation coefficients are not available.

Good predictors, overall, were "days since cropping," "cumulative rainfall since cropping," and "hours since start of storm." The fact that "days since cropping" was a good predictor variable would tend to make "cumulative rainfall since cropping" a good predictor as well (the correlation coefficient between the two predictors was typically about 0.98, making these variables virtually identical in terms of predictive value). "Hours since start of storm" was also a fair predictor, especially for atrazine and nitrate. However, it too was highly correlated ($r = 0.74$) with "days since cropping," suggesting that the length of storms (or perhaps sampling times within storms) depended upon time of the year.

A comment is in order here regarding the high correlations between atrazine concentration and the various predictor variables. In this data base, few satisfactory observations for atrazine concentrations were available (about 20). Accordingly, these high correlations must be viewed with suspicion, as they may well be spurious--more a result of special conditions than

TABLE 23. SUMMARY LISTING OF POTENCY FACTORS ESTIMATED FROM AGRICULTURAL, SUBURBAN, AND URBAN WATERSHEDS EXAMINED IN THIS STUDY. SEE NOTE A

	Watkinsville, Georgia Plot P-04	Watkinsville, Georgia Plot P-03	Watkinsville, Georgia Plot P-01	Buffalo Bill, Iowa Watershed	Michigan State University Test Plots	Viewridge 1, Seattle	South Seattle	Southcenter, Seattle	Honey Creek, Melmore, Ohio
Atrazine, ppb	1.02 .105 (B,C,D,E)				1.56 .298 (B,F)				
TKN, ppm	1.30 (B)			3.6×10^{-4} (B)	.271 (C,D)				2.1×10^{-3} (B)
NH_4-N, ppm	.850			6.3×10^{-5} (B)	.164 (B)	-2.05 .336 (B,F)			-.745 -.155 (B,F)
NO_3-N, ppm	-.821 .541 (B,E)			4.7×10^{-6} (B,C,D)	-.020 (B,C,D)				
PO_4-P, ppm	.062 (B)			2.1×10^{-5} (B)		-1.43 .172 (B,F)			-1.15 .021 (B,C,D,F)
Cl, ppm	.168 .247 (B,E)								1.77 -.172 (B,F)
Trifluralin, ppb		.694 (B)	.579 .0016 (B,E)						
Diphenamid, ppb		62.7	.277 .023 (B,E)						
Dieldrin, ppt				.040					
DDT, ppt				.0002 (B,C,D)					
DDT, ppt				.0003 (B,C,D)					
NO_2-N, ppm				1.5×10^{-5} (B)					
Fecal Coliforms, MPN/100ml				4.15 1.7×10^{-4} (B,E)					
Reference Page Number	52	53	54	58	60	61	66	69	72

TABLE 23 (continued)

	Watkinsville, Georgia Plot P-04	Watkinsville, Georgia Plot P-03	Watkinsville, Georgia Plot P-01	Buffalo Bill, Iowa Watershed	Michigan State University Test Plots	Viewridge 1, Seattle	South Seattle	Southcenter, Seattle	Honey Creek, Melmore, Ohio
BCO, ppm				1.7×10^{-3} (B)					
Total P, ppm					.053 (B)	-.917 .237 (B,F)	-1.40 .419 (B,F)	-.841 .190 (B,F)	1.1×10^{-3} (B)
$NO_2 + NO_3$-N, ppm						-.108 -.140 (F)	-.967 .252 (B,F)	-.760 .375 (B,F)	.266 .214 (B,F)
Organic N, ppm						-.447 .280 (B,F)	-.567 .211 (B,F)	-.316 .153 (B,F)	
Conductivity µmhos/cm						144. -24.6 (B,E)			2.94 -.152 (B,G)
Turbidity, JTU						.537 .358 (B,F)			
Lead, ppb						1.42 4.90 (B,F)	1.42 .497 (B,F)	.210 .179 (B,F)	
Zinc, ppb						1.47 .283 (B,F)	1.91 .256 (B,F)	2.24 .137 (B,F)	
Iron, ppb									.040 (B)
Total P-OPO_4-P, ppm									1.1×10^{-2} (B)
SiO_2, ppm									.679 .076 (B,F)
Reference Page Number	52	53	54	58	60	61	66	69	72

Notes: A. Potency factors (m from $C = b + mS$) provided in this table are for untransformed data **unless otherwise** noted. Commonly, the transformed data proved to be much more reliable than untransformed in terms of statistical results, as explained in the text. Where transformations are noted, regression results for untransformed data were judged unreliable. Note that C = concentration of pollutant, S = concentration of suspended solids. For use with NPS, except at Watkinsville and Michigan State U. Farms, multiply m by 100%. Where constituent C is in ppm, result is in grams C per gram of S. Where C is in ppb, result is in milligrams C per gram of S. For ppt, result is micrograms C per gram of S. For Watkinsville and Michigan data, instead of multiplying by 100%, divide by 10 and interpret as just described. For example, for ammonia at Watkinsville Plot P-04, we have .85 milligrams of ammonia per gram of sediment (because sediment is in g/ℓ) and the potency factor is .85/10 = .085 (= .00085 gram per gram of sediment). For **dieldrin (Buffalo Bill shed), the potency factor is .04 ppt x 100% = 4 (= .04 micrograms per gram of sediment).**

B. Intercept (b) statistically non-zero at 95% confidence level in a two-tailed test, but ignored here. If B is not specified, intercept was not significant.

C. Overall regression not significant at 95% confidence level.

D. Potency factor (m) not significant at 95% confidence level.

E. $\text{Log}_{10}\, C = b + mS$. Upper value is b, lower is m.

F. $\text{Log}_{10}\, C = b + m \log_{10} S$. Upper value is b, lower is m.

G. $C = b + m \log_{10} S$. Upper value is b, lower is m.

TABLE 24. VARIABLES EXAMINED IN MULTIPLE REGRESSION AS CANDIDATE PREDICTORS OF RUNOFF WATER QUALITY

- Elapsed time since reference date
- Month
- Elapsed time since storm began
- Dry days preceding storm (antecedent dry days)
- Suspended solids concentration
- Water quality variables
- Flow
- Storm rainfall

TABLE 25. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT AND INDEPENDENT VARIABLES FOR WATKINSVILLE PLOT P-04 DATA

	Hours since start of storm	Days since cropping	Cumulative rain since cropping, inches	Cumulative rainfall during storm, inches	Rainfall intensity at time (t-1) hours	Rainfall intensity at sampling time (t)	Suspended sediment concentration	Atrazine concentration in sediment	Ammonia concentration in runoff	TKN concentration in runoff	Total phosphorus concentration in runoff	Ammonia concentration in sediment	Nitrate concentration in runoff	TKN concentration in sediment	Dissolved phosphate concentration	Total phosphorus concentration in sediment	Dissolved chloride concentration
Dissolved atrazine*	.971	-.979	-.965	.824	.382	-.099	.078	.868*	-.601	-.755	.405						
Dissolved TKN	-.384	-.492	-.418	-.424			.609		.786*	1.00	-.033	-.267*	.681*	-.170*	.071	-.265*	.236*
Dissolved ammonia	-.262	-.655	-.601	-.026	.014	-.032	.430		1.00	.783	-.025	-.099	.697	-.138	.048	-.106	.402
Dissolved nitrate	-.592	-.771	-.754	-.191			.444		.754		-.415	.061*	1.00	-.077*	-.269	-.188*	.404
Dissolved Phosphate	.347	.243	.320	-.225			.264		.084*	.130*		-.171*	-.269*	-.062*	1.00	-.194*	-.031*
Dissolved chloride	-.313	-.595	-.579	.282	.246*	.366*	.236		.354*	.391	.026	.015*	.579*	-.433*	.133	-.267*	1.00

*Variables transformed by $\log_{10}$.

of any real, persistent, physical relationship. For the remainder of dependent variables, there were typically about 66 observations, making correlations somewhat more reliable.

Suspended sediment concentration generally proved a modest predictor with correlations (r) ranging from 0.444 down to 0.078 except for dissolved TKN (r = 0.609). In all cases except TKN, the r^2 value shows that suspended solids concentration could account for less than 20% of the variance of the water quality constituent. Thus about 80% of the variance of dissolved nitrate cannot be accounted for by the variance of suspended solids. Only 46% of dissolved TKN variance may be attributed to variance in suspended solids.

With the exception of dissolved phosphate, all other water quality variables correlated fairly well with one or more predictor variables. In addition to suspended solids, TKN correlated fairly well with "days since cropping" (almost 25% of TKN variance), with runoff ammonia, and with runoff nitrate (almost 50% of variance in common). Dissolved ammonia correlated well with "days since cropping" and with nitrate.

Table 26 presents multiple regression results for Watkinsville plot P-04. The most important non-water quality variables were "hours since start of storm," and "cumulative storm rainfall." It is of interest to note that in some regressions (e.g., for dissolved TKN) "hours since start of storm" correlated more strongly with the dependent variable than did "cumulative storm rainfall," suggesting a number of possible phenomena. These would include volatilization, biodegradation, or photodegradation at the surface, coupled with initial resistance to transport of sediment and/or pollutant, perhaps as a result of slow soil wetting. In further support of this, there were cases (e.g., dissolved TKN and phosphate as dependent variables) where, in later regression steps than that shown in Table 26, both "cumulative storm rainfall" and "time since start of storm" entered the regression.

Table 27 provides regression equations which may be of use for comparing results of other studies, or as input in designing future investigations.

Watkinsville: Diphenamid and Trifluralin Residues

Data from Watkinsville plots P-01 and P-03 included runoff diphenamid and trifluralin concentrations. In contrast to the previous discussion (plot P-04 data), analyses here encompassed more than 1 year of data. Data were pooled so that for P-01, 2 years (1972 and 1973) were represented, and for P-03, 4 years (1972-1975) were represented.

Table 28 shows correlations (r) between diphenamid and trifluralin data and the nine predictor variables. As shown, four variables were strongly correlated with dissolved trifluralin and diphenamid in the runoff. These were "days since cropping," "cumulative rainfall since cropping," "suspended sediment," and herbicide (diphenamid or trifluralin) concentration in the sediment. The fact that time and rainfall since cropping strongly correlated with the dependent variables suggests that available herbicide decreases (or is less available for washoff) over time. These two predictor variables ("time" and "rainfall since cropping") had 83% of their variance in common, and

TABLE 26. MULTIPLE REGRESSION STATISTICS FOR WATKINSVILLE PLOT P-04. ENTRIES ARE t VALUES FOR EACH REGRESSION COEFFICIENT

	Dissolved atrazine, mg/ℓ †*	TKN, mg/ℓ	Ammonia N, mg/ℓ†	Nitrate N, mg/ℓ	Phosphate P, mg/ℓ†	Chloride, mg/ℓ†
Hours since start of storm	13.5		-5.12		3.92	
Days since cropping						-6.45
Cumulative rainfall since cropping, inches						
Cumulative rainfall during storm, inches		-3.00				2.55
Rainfall intensity at t-1, in/hr**						*
Rainfall intensity at t, in/hr						*
Suspended solids concentration, g/ℓ						
Sediment atrazine, μg/ℓ	*					
Dissolved ammonia N, mg/ℓ		9.58*		7.93	*	*
Dissolved TKN, mg/ℓ			9.40		*	
Dissolved total phosphorus, mg/ℓ						
Sediment ammonia N concentration, mg/ℓ		*			3.32*	*
Dissolved nitrate N concentration, mg/ℓ		*			*	*
Sed. TKN concentration, mg/ℓ		*			*	-3.79*
Dissolved phosphate P concentration, mg/ℓ						
Sediment total phosphorus concentration, mg/ℓ		*			*	*
Dissolved chloride concentration, mg/ℓ		*		4.45	*	
R^2	.816	.666	.690	.629	.252	.465
n	42	66	108	66	66	108
$n^{\dagger\dagger}$	8	47	12	66	13	25

*Variables $\log_{10}$ transformed

**t-1 implies runoff sampling time minus one hour.

†Substantial autocorrelation of residuals was noted. Estimated number of independent observations very much less than n. Estimated $n^{\dagger\dagger}$ based on an approximation due to Wold (see Appendix A).

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at $\alpha = .05$, two-tailed test.

TABLE 27. REGRESSION EQUATIONS WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSIONS ONLY, α = .05, TWO-TAILED TEST) FROM WATKINSVILLE PLOT P-04

Equations	R^2
$\log_{10}[A] = .405 \pm .12 + (.00478 \pm .0006)\ \mathrm{HSS}$	.943
$\log_{10}[A] = 9.13 \pm .79 - (.047 \pm .0046)\ D$	.960
$\log_{10}[A] = 3.49 \pm .32 - (.311 \pm .04)\ \mathrm{CR}$	.931
$[A] = (2.14 \times 10^{-5} \overset{\times}{\div} .025)\ [A_{sed}]^{(2.44 \pm .67)}$	.754
$[\mathrm{TKN}_D] = 4.12 \pm .92 - (.0082 \pm .0037)\ D$	.242
$[\mathrm{TKN}_D] = 2.64 \pm .23 + (1.19 \pm .317) \log_{10}[\mathrm{NO}_3]$	.464
$[\mathrm{TKN}_D] = 2.80 \pm .21 + (1.62 \pm .32) \log_{10}[\mathrm{NH}_{4_D}]$	.619
$[\mathrm{TKN}_D] = 3.13 - .213\ \mathrm{CR} + 1.50 \log_{10}[\mathrm{NH}_{4_D}]$	.666
$[\mathrm{NH}_{4_D}] = 2.96 \pm .56 - (.0094 \pm .002)\ D$	.429
$[\mathrm{NH}_{4_D}] = 1.50 \pm .29 - (.035 \pm .014)\ \mathrm{CR}$	.361
$[\mathrm{NH}_{4_D}] = 1.80 \pm .126 + (.480 \pm .095)\ [\mathrm{NO}_3]$	.486
$[\mathrm{NH}_{4_D}] = -.745 \pm .212 + (.661 \pm .101)\ [\mathrm{TKN}_D]$	.613
$[\mathrm{NH}_{4_D}] = .779 - .0047\ D + .512\ [\mathrm{TKN}_D]$	.690
$[\mathrm{NO}_3] = .109 + .898\ [\mathrm{NH}_{4_D}] + .083\ [\mathrm{Cl}]$	.630
$[\mathrm{PO}_4] = .053 + .0001\ \mathrm{HSS} + .088\ [\mathrm{SS}]$	.252
$\log_{10}[\mathrm{Cl}] = 1.53 \pm .32 - (.0045 \pm .0012)\ D$	.354
$\log_{10}[\mathrm{Cl}] = .852 \pm .155 - (.018 \pm .0049)\ \mathrm{CR}$	.335
$[\mathrm{Cl}] = (3.71 \overset{\times}{\div} 1.24)\ ([\mathrm{NO}_3] + 1.)^{(.585 \pm .158)}$*	.335
$\log_{10}[\mathrm{Cl}] = 1.79 - .0037\ D + .084\ \mathrm{CRS} - .190 \log_{10}([\mathrm{TKN}_S] + 1.)$*	.465

Explanation of symbols on following page.

TABLE 27 (continued)

Explanation of symbols:

A = dissolved atrazine, ppb

HSS = hours since start of storm

D = days since cropping

CR = cumulative rainfall since cropping, in.

A_{sed} = sediment atrazine, ppb

TKN_D = dissolved total kjeldahl nitrogen, ppm

NH_{4_D} = dissolved ammonia nitrogen, ppm

NO_3 = dissolved nitrate nitrogen, ppm

Cl = dissolved chloride, ppm

PO_4 = dissolved phosphate phosphorus, ppm

SS = suspended solids, g/ℓ

TKN = sediment total kjeldahl nitrogen, ppm

---

*A constant (small relative to the values of the variables themselves) was added to independent variable to prevent $\log_{10} 0$.

TABLE 28. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT VARIABLES (TRIFLURALIN AND DIPHENAMID) AND INDEPENDENT VARIABLES DATA ARE FROM WATKINSVILLE TEST PLOTS

	Elapsed Time†	Days since cropping	Cumulative rain since cropping	Cumulative storm rainfall	Rain intensity at (t-1)* hr**	Rain intensity at (t-0) hr	Suspended sediment	Sediment diphenamid concentration	Sediment trifluralin concentration
Trifluralin (P-01)	-.180	-.812	-.771	-.371	-.366	.002	.665		.141
Diphenamid (P-01)	.126	-.758	-.810	-.254	-.211	-.227	.666	.713	
Trifluralin (P-03)	-.561**	-.760**	-.884**	-.187	-.100**	.025**	.538		.633**
Diphenamid (P-03)	-.377**	-.825**	-.870	-.582	-.214**	-.324**	.615**	.513	

*(t-1) implies rainfall data at one hour before runoff was sampled.

All variables marked "" are $\log_{10}$ transformed. Data for P-01 are for 1972 and 1973. Data for P-03 are for 1972-1975.

†Elapsed time is days since the first data collection. Thus, this variable should contain information on long-term changes in runoff such as caused by monotonic increases or decreases in amount of herbicide applied each year or monotonic climatologic trends.

partialing "cumulative rainfall since cropping" out of "time since cropping" and then correlating the latter with $\log_{10}$ trifluralin (plot P-03) reduced the shared variance between predictor and dependent variables by 77% (from 78% to 0.9%). Accordingly, as much as 98% of the variance of plot P-03 trifluralin shared with "time since cropping" might be caused by storm washoff. It is possible, of course, that less is directly caused by washoff, since such effects as wind erosion losses and natural degradation would also be functions of time and would correlate with both time and cumulative rainfall.

The correlations with suspended solids and sediment herbicide concentrations have been discussed earlier.

Table 29 presents t statistics for multiple regression results with trifluralin and diphenamid as the dependent variables. Table 30 presents regression equations for this data set.

Buffalo Bill Watershed

Table 31 presents correlations (r) between dependent and independent variables for the Buffalo Bill Watershed. As shown, there are several variables which correlate very poorly with virtually all others (i.e., phosphate, DDT, DDE and nitrate) while some (i.e., fecal coliforms. BOD, dieldrin) correlate moderately well with a number of other variables. Dieldrin, for example, shares 47, 49, and 32% of its variance with suspended solids, turbidity, and fecal coliforms, respectively (based on r^2). For BOD correlations with suspended solids, turbidity, and fecal coliforms, the values are 32, 30, and 40%, respectively. In contrast, nitrate correlates most strongly with nitrite (22% shared variance) and ammonia (11% shared variance) and phosphate most strongly correlates with TKN, sharing only 10% common variance. In the case of BOD, dieldrin, and fecal coliforms, some commonality is suggested in the mechanisms and/or rates of deposition, degradation, and runoff; while for those not correlating well, other factors not accounted for by the available data must be involved.

Table 32 presents multiple regression statistics for the Buffalo Bill Watershed data base. As shown in the table, even in the best regressions (highest R^2 values), only slightly over 50% of the variance of the dependent variables could be explained, and in many cases R^2 was substantially lower. Unfortunately, only water quality variables were available in this data base. It is likely that land use, farming practice, and storm information would improve explication.substantially.

Table 33 presents regression equations for the Buffalo Bill Watershed data base.

Michigan State University Test Plots

Because of uncertainties and errors found in the data, multiple regression results (like results from this data base presented earlier) are suspect. Accordingly, multiple regression results will not be presented.

TABLE 29. MULTIPLE REGRESSION STATISTICS FOR WATKINSVILLE PLOTS P-01 AND P-03 DATA. ENTRIES ARE VALUES OF t FOR EACH REGRESSION COEFFICIENT

	Dissolved trifluralin concentration, ppb plot P-01	Dissolved diphenamid concentration, ppb plot P-01	Dissolved trifluralin concentration, ppb plot P-03	Dissolved diphenamid concentration, ppb plot P-03
Elapsed time	3.99 *	*	-6.47	2.05
Days since cropping	-13.1 *	*		-11.9
Cumulative rain since cropping, in.	*	*	-24.3	-10.6 *
Cumulative storm rainfall, in.	*	*	8.29 *	-7.27 *
Rain intensity at (t-1)**, in./hr.	*	*		
Rain intensity at (t-0), in./hr.	*	-2.12 *		-2.16
Suspended sediment, g/ℓ	6.08 *	12.2 *	*	
Sed. diphenamid concentration, ppb		*		5.87 *
Sed. trifluralin concentration, ppb	*		5.12	
R^2	.762	.382	.892	.898
n	134	260	170	223
$n^{††}$				

*Variables not $\log_{10}$ transformed.

**(t-1) implies runoff sampling time minus one hour.

††Autocorrelation statistics not available for these regressions. See footnote of Table 28.

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at $\alpha = .05$, two-tailed test.

TABLE 30. REGRESSION EQUATIONS WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSIONS ONLY, α = .05, TWO-TAILED TEST) FROM WATKINSVILLE PLOTS P-01 AND P-03 HERBICIDE DATA

Equations*	R^2
Plot P-01	
$\log_{10}([\mathrm{Trif}] + 1.) = .937 \pm .035 - (.0082 \pm .001)\ D$	.659
$\log_{10}([\mathrm{Trif}] + 1.) = .954 \pm .046 - (.099 \pm .015)\ RSC$	.594
$\log_{10}([\mathrm{Trif}] + 1.) = .735 + 4.6 \times 10^{-4}\ ET - 7.33 \times 10^{-3}\ D + 9.1 \times 10^{-3}\ SS$	.762
$\log_{10}([\mathrm{Diph}] + 1.) = 2.21 \pm .07 - (.024 \pm .0025)\ D$	.574
$\log_{10}([\mathrm{Diph}] + 1.) = 2.12 \pm .07 - (.181 \pm .016)\ RSC$	.655
$\log_{10}([\mathrm{Diph}] + 1.) = .595 \pm .075 + (1.45 \times 10^{-3} \pm 1.8 \times 10^{-4})\ [\mathrm{Diph}_{sed}]$	.509
Plot P-03	
$([\mathrm{Diph}] + 1.) = (1175 \overset{\times}{\div} 1.38)\ (D + 1.)^{(-1.12 \pm .11)}$	.681
$\log_{10}([\mathrm{Diph}] + 1.) = 2.69 \pm 0.1 - (.246 \pm .018)\ RSC$	.756
$\log_{10}([\mathrm{Diph}] + 1.) = 2.42 \pm 0.18 - (.759 \pm .141)\ RSS$	.338
$\log_{10}([\mathrm{Diph}] + 1.) = 1.03 \pm .164 + (7.88 \times 10^{-4} \pm 1.75 \times 10^{-4})\ [\mathrm{Diph}_{sed}]$	.263
$\log_{10}([\mathrm{Diph}] + 1.) = 3.20 - .515 \log_{10}(D + 1.) - .139\ RSC - .305\ RSS$	.874
$([\mathrm{Trif}] + 1.) = (32.4 \overset{\times}{\div} 1.51)\ (ET + 1.)^{(-.443 \pm .100)}$	.315
$([\mathrm{Trif}] + 1.) = (14.8 \overset{\times}{\div} 1.17)\ (D + 1.)^{(-.357 \pm .046)}$	.577
$([\mathrm{Trif}] + 1.) = (13.5 \overset{\times}{\div} 1.10)\ (RSC + 0.1)^{(-.853 \pm .069)}$	.782
$([\mathrm{Trif}] + 1.) = (1.56 \overset{\times}{\div} 1.29)\ ([\mathrm{Trif}_{sed}] + 1.)^{(.315 \pm .059)}$	.400
$\log_{10}([\mathrm{Trif}] + 1.) = .824 - .823 \log_{10}(RSC + 0.1) + .102\ RSS + .114 \log_{10}([\mathrm{Trif}_{sed}] + 1.)$	.864

*Note: Constants (small relative to values of the variables themselves) were added to prevent $\log_{10} 0$.

TABLE 30 (continued)

Explanation of symbols:

Diph = dissolved diphenamid, ppb

D = days since cropping

RSC = cumulative rain since cropping, inches

RSS = cumulative storm rainfall, inches

$\mathrm{Diph}_{\mathrm{sed}}$ = sediment diphenamid, ppb

Trif = dissolved trifluralin, ppb

ET = elapsed time since 7/02/72, days

$\mathrm{Trif}_{\mathrm{sed}}$ = sediment trifluralin, ppb

SS = suspended solids, g/ℓ

TABLE 31. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT AND INDEPENDENT VARIABLES FOR THE BUFFALO BILL WATERSHED

	Suspended solids concentration	Turbidity	BOD concentration	Fecal coliforms	TKN concentration	Ammonia concentration	Nitrite concentration	Nitrate concentration	Phosphate concentration
Dieldrin*	.685	.697	.482	.569*	.317	.300	.282*	.088*	.020*
DDT*	.195	.181	.185	.230*	-.055	.207	.362*	-.235*	.084*
DDE	.143	.074	.132	.226*	-.004	-.074	-.024	-.124	.015
Phosphate*	-.289	-.249	.059	.159*	.324	-.049	.178*	.195*	1.00*
Nitrate*	.016	.115	.124	.115*	.113	.330	.425*	1.00*	.195*
Nitrite*	.325	.473	.375	.285	.316	.632	1.00*	.469	.178*
Ammonia	.387	.508	.339	.197*	.207	1.00	.632*	.332	-.049*
TKN	.339	.392	.438	.393*	1.00	.207	.316*	.193	.324*
Fecal coliforms*	.463	.502	.632	1.00*	.393	.335	.373	.144	.230
BOD	.566	.545	1.00	.632*	.438	.339	.375*	.094	.059*

*Variables $\log_{10}$ transformed.

TABLE 32. MULTIPLE REGRESSION STATISTICS FOR THE BUFFALO BILL WATERSHED. ENTRIES ARE VALUES OF t FOR EACH REGRESSION COEFFICIENT

	Dieldrin concentration, ppt*†	DDT concentration, ppt	DDE concentration, ppt	Phosphate concentration, mg/ℓ*	Nitrate N concentration, mg/ℓ*	Nitrite N concentration, mg/ℓ*†	Ammonia N concentration, mg/ℓ	TKN concentration, mg/ℓ	Fecal coliforms, MPN/100 ml*	BOD concentration, mg/ℓ
Suspended solids concentration, mg/ℓ				-5.29						3.65
Turbidity, JTU	6.61					2.61	2.99			
BOD concentration, mg/ℓ								4.60	6.89	
Fecal coliforms, MPN/100 ml	3.68 *	*	2.09 *	2.38 *	*			*		4.74 *
TKN concentration, mg/ℓ				4.12						2.02
Ammonia-N concentration, mg/ℓ						4.46				
Nitrite-N concentration, mg/ℓ	*	2.45 *		*	4.38 *		5.60 *	*	2.67	*
Nitrate-N concentration, mg/ℓ	*	*		*		3.82				
Phosphate concentration, mg/ℓ	*			*	*	*	*	3.27 *		*
R^2	.559	.131	.051	.328	.181	.514	.456	.281	.446	.518
n	85	42	83	89	89	89	89	89	89	89
$n^{\dagger\dagger}$	14	42	61	89	53	35	42	52	89	89

*Variables not $\log_{10}$ transformed.

†Substantial autocorrelation of residuals was noted. Estimated number of independent observations very much less than n.

††Estimated n, based upon an estimation due to Wold (see Appendix A) shown as $n^{\dagger\dagger}$.

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at α = .05, two-tailed test.

TABLE 33. REGRESSION EQUATIONS WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSION ONLY, $\alpha = .05$, TWO-TAILED TEST) FROM THE BUFFALO BILL WATERSHED DATA BASE

Equations	R^2
$\log_{10} [\mathrm{Diel}] = .913 \pm .091 + (6.11 \times 10^{-4} \pm 1.37 \times 10^{-4})\ T$	.486
$\log_{10} [\mathrm{Diel}] = .889 \pm .149 + (.030 \pm .012)\ [\mathrm{BOD}]$	.232
$[\mathrm{Diel}] = (.658 \substack{\times \\ \div} 2.79)\ \mathrm{FC}^{(.309 \pm .098)}$	.323
$\log_{10} [\mathrm{Diel}] = 4.8 \times 10^{-4}\ T + .167 \log_{10} \mathrm{FC}$	.486
$[\mathrm{DDE}] = 1.91 + 1.01 \log_{10} \mathrm{FC}$	.051
$\log_{10} [\mathrm{DDT}] = 1.02 + .276 \log_{10} [\mathrm{NO_2}] - .361 \log_{10} [\mathrm{NO_3}]$	.250
$\log_{10} [\mathrm{PO_4}] = -.910 - 8.0 \times 10^{-5}\ \mathrm{SS} + .074\ [\mathrm{TKN}]$	.283
$\log_{10} [\mathrm{NO_3}] = .433 + .316 \log_{10} [\mathrm{NO_2}]$	.181
$\log_{10} [\mathrm{NO_2}] = -1.38 \pm .09 + (3.43 \times 10^{-4} \pm 1.36 \times 10^{-4})\ T$	.223
$\log_{10} [\mathrm{NO_2}] = -1.49 \pm .08 + (.585 \pm .153)\ [\mathrm{NH_4}]$	.400
$\log_{10} [\mathrm{NO_2}] = -1.48 \pm .12 + (.198 \pm .079)\ [\mathrm{NO_3}]$	.220
$\log_{10} [\mathrm{NO_2}] = -1.61 + .495\ [\mathrm{NH_4}] + .123\ [\mathrm{NO_3}]$	.476
$[\mathrm{NH_4}] = .261 \pm .093 + (3.99 \times 10^{-4} \pm 1.44 \times 10^{-4})\ T$	.259
$[\mathrm{NH_4}] = 1.28 \pm .22 + (.684 \pm .179) \log_{10} [\mathrm{NO_2}]$	.400
$[\mathrm{NH_4}] = 1.02 + 2.0 \times 10^{-4}\ T + .546 \log_{10} [\mathrm{NO_2}]$	.456
$[\mathrm{TKN}] = 3.39 + .154\ [\mathrm{BOD}] + 1.92 \log_{10} [\mathrm{PO_4}]$	.281
$\log_{10} \mathrm{FC} = 4.15 \pm .20 + (1.66 \times 10^{-4} \pm 6.77 \times 10^{-5})\ \mathrm{SS}$	.214
$\log_{10} \mathrm{FC} = 4.05 \pm .20 + (8.62 \times 10^{-4} \pm 3.17 \times 10^{-4})\ T$	.252
$\log_{10} \mathrm{FC} = 3.67 \pm .25 + (7.65 \times 10^{-2} \pm .02)\ [\mathrm{BOD}]$	.400

TABLE 33 (continued)

Equations	R^2
$\log_{10} FC = 3.58 + .069 [BOD] + 1.94 [NO_2]$	.446
$[BOD] = 6.71 \pm 1.65 + (7.75 \times 10^{-3} \pm 2.54 \times 10^{-3}) T$	.297
$[BOD] = -12.9 \pm 6.2 + (5.22 \pm 1.36) \log_{10} FC$	.400
$[BOD] = -8.96 + .001 SS + 3.89 \log_{10} FC$	.495

Explanation of symbols:

Diel = dissolved dieldrin, ppt

T = turbidity, JTU

BOD = biochemical oxygen demand (not specified whether 5-day, ultimate or other), ppm

FC = fecal coliforms, MPN/100 ml

DDE = dissolved DDE, ppt

DDT = dissolved DDT, ppt

NO_2 = dissolved nitrite-N, ppm

NO_3 = dissolved nitrate-N, ppm

TKN = dissolved total kjeldahl nitrogen, ppm

NH_4 = dissolved ammonia-N, ppm

PO_4 = dissolved phosphate, ppm (unknown whether as P or as PO_4)

SS = suspended solids, ppm

Seattle: Southcenter

Tables 34 and 35 show multiple regression results for Southcenter. Table 34 provides simple linear correlations between dependent and independent variables. Table 35 shows multiple regression statistics.

In Table 34, it can be seen that in all cases, there are better predictors of runoff water quality than suspended solids, although in general, such data may not be as readily available or as practical from a standpoint of modeling applications. In the case of lead, organic nitrogen has a higher r value. The value of r^2 for lead versus suspended solids is 0.446 ($r = 0.668$) while r^2 for organic nitrogen is 0.503, which is probably significantly better. Zinc concentration is much more strongly correlated with lead concentration than with suspended solids, suggesting common phenomena other than association with suspended matter. The common phenomena may include those related to zinc and lead deposition, and these may not be the same as those involved in dust and dirt deposition. The upper accumulation limits or rate of approach to those limits for particulate and for lead and zinc may be very different. Further, washoff characteristics for lead and zinc may also differ from those of suspended solids.

Both lead and zinc correlate strongly with nitrite plus nitrate and with organic nitrogen. They also correlate well with total phosphorus.

Urban runoff water quality models, such as SWMM, model water quality as a simple linear function of runoff suspended sediment, the accumulation and availability of which are, in turn, simulated as a function of the number of dry days preceding each storm. Lead and zinc concentrations are only weakly correlated with the number of dry days, suggesting that "antecedent dry days" is not a very good predictor variable for lead and zinc. In this data base, suspended solids shares only 2% ($r^2 = 0.02$) of its variance with the number of antecedent dry days, making dry days a poor predictor of suspended solids as well.

Why the correlation is so poor is not clear since it would be expected that runoff lead and zinc concentration would be proportional to the amount available for transport. Since washoff should, in turn, affect amount available to wash off, it is reasonable to expect runoff lead and zinc concentrations to be a function of dry days preceding storms.

Total phosphorus and organic nitrogen similarly correlate weakly with "antecedent dry days." In contrast, the relatively soluble species, nitrite plus nitrate nitrogen, orthophosphate phosphorus, and ammonia nitrogen all are much better predicted by "antecedent dry days" than by suspended solids concentration. Since the latter is itself poorly correlated with "antecedent dry days," it seems that simulating the species nitrate, nitrite, orthophosphate, and ammonia as a function of dry days preceding storms through the intermediary suspended solids is very tenuous. It should be noted here, however, that the poor correlation between "antecedent dry days" and suspended solids may well be due to dilution effects related to variations in flow. "Antecedent dry days" may, in fact, correlate well with dust and dirt accumulation. Nonetheless, the poor correlation between sediment and some water quality

TABLE 34. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT AND INDEPENDENT VARIABLES FOR SOUTHCENTER, SEATTLE

	Days since 1/1/73*	Month*	Hours since storm start	Dry days before storm	Suspended solids	Lead concentration	Zinc concentration	Nitrite plus nitrate N concentration	Ammonia N concentration	Total phosphorus concentration	Orthophosphate phosphorus concentration	Organic nitrogen concentration	Conductivity	Turbidity	Flow	Storm rainfall (cumulative)*
Lead	.084	.186	-.282	.231	.668	1.00	.666	.602	.351	.486	.381	.709	.151	.548	-.073	-.273
Zinc	.203	-.040	-.202	.140	.374	.666	1.00	.700	.331	.520	.427	.654	.495	.489	-.266	-.311
Nitrite plus Nitrate N	-.014	.229	-.131	.483	.298	.602	.700	1.00	.583	.472	.614	.662	.420	.271	-.273	-.285
Ammonia N	-.251	.374	.106	.523	.135	.351	.331	.583	1.00	.207	.546	.286	.062	.075	-.128	-.211
Total Phosphorus	-.130	-.161	-.292	.138	.390	.486	.520	.472	.207	1.00	.540	.532	.534	.568	-.273	-.230
Orthophosphate Phosphorus	.113	.401	.070	.402	.112	.381	.427	.614	.546	.540	1.00	.481	.391	.346	-.277	-.308
Organic Nitrogen	.225	.267	-.287	.265	.514	.709	.654	.662	.286	.532	.481	1.00	.316	.471	-.153	-.239

All variables but those marked with "" are $\log_{10}$ transformed.

TABLE 35. MULTIPLE REGRESSION STATISTICS FOR SOUTHCENTER, SEATTLE. ENTRIES ARE VALUES OF t FOR EACH REGRESSION COEFFICIENT

	Lead concentration, μg/ℓ	Zinc concentration, μg/ℓ	Nitrite and nitrate N concentration, mg/ℓ†	Organic nitrogen concentration, mg/ℓ	Total phosphorus concentration, mg/ℓ	
Days since 1/1/73		8.40		6.29*	-3.94	Long-term trends
Month of the year	*	-8.67*	*		-3.42*	Storm-related variables
Hours since start of storm	*		-3.06		*	
Cumulative storm rainfall, inches	-4.85*	*	*	-2.51*	*	
Dry days before storm			5.50			
Suspended solids concentration, mg/ℓ	8.52			3.79		Suspended matter
Turbidity, JTU		4.23			7.63*	
Lead concentration, μg/ℓ		4.64		2.49		Heavy metals
Zinc concentration, μg/ℓ	4.83		7.96			
Nitrate and nitrite N concentration, mg/ℓ	2.61	9.32		6.66		Nitrogen variables
Ammonia N concentration, mg/ℓ		3.95	3.90			
Organic nitrogen concentration, mg/ℓ	3.80		2.83		5.37	
Total phosphorus concentration, mg/ℓ				3.77		Phosphorus variables
Orthophosphate P concentration, mg/ℓ			3.41			
Conductivity, μmhos/cm	-3.61		*		6.82*	
Flow, cfs	*				*	
R^2	.738	.751	.744	.684	.634	
n	199	199	199	199	199	
$n^{††}$	99	98	32	199	199	

*Variables not $\log_{10}$ transformed.

†Substantial autocorrelation of residuals was noted. Estimated number of independent observations very much less than n (199), probably equivalent to about 32 based upon an estimation due to Wold (see Appendix A). Estimated n shown as $n^{††}$.

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at α = .05, two-tailed test.

constituents and the suggested low correlation between "antecedent dry days" and dust and dirt accumulation seem to warrant a dry days - water quality estimation procedure.

Regarding other predictors, "elapsed time," "hours since start of storm," "runoff flow," and "cumulative storm rainfall" were not well correlated with any of the dependent variables. "Month of the year" was modestly correlated with soluble species, perhaps suggesting the effect of seasonality of "antecedent dry days" or of deposition phenomena. As might be expected, turbidity was correlated with the dependent variables to an extent similar to suspended solids. Conductivity correlated well with some dependent variables.

Multiple regression results shown in Table 35 suggest that phenomena influencing runoff water quality with respect to the dependent variables shown, are quite complex. This is suggested by the fact that different predictor variables were selected for the several dependent variables. Further, and perhaps more importantly, no single (or small number of) independent variables was adequate to explain water quality variability. If simple phenomena were involved, it would be expected that perhaps two or three predictors would account for much of the variability of runoff water quality. It is possible, of course, that the set of predictors used here simply missed one or two critical ones. It is of interest to note that if the predictor variables are grouped as shown in Table 35 (e.g., into storm, suspended matter, heavy metal, nitrogen, and phosphorus variables), then some consistency appears in the set of predictor variables chosen for each dependent variable. In all five regressions, at least one storm variable and one nitrogen variable were selected. In four of the five regressions, at least one heavy metal, and one suspended solids variable was selected. This suggests that although different nitrogen variables were chosen in different regressions, it may well have been possible to obtain nearly as high values of R^2 using only one candidate variable in each variable category. This suggests, further, that the concentration of any of the water quality variables in runoff depends upon phenomena reflected in information contained in perhaps four general variables (i.e., storm characteristic, heavy metal, suspended solids, nitrogen).

The suggested complexity of phenomena influencing runoff water quality apparently does not preclude some spatial consistency. When regression coefficients from Viewridge 1 (Seattle) were used in conjunction with Southcenter data, the equations from Viewridge 1 were, in some cases, reasonably good predictors of runoff quality at Southcenter. Table 36 shows "validation-generalization" results.

Table 37 presents predictive equations which may be of use in estimating runoff water quality from urban watersheds, or which may be of interest in comparing with results of other studies.

Seattle: South Seattle

Table 38 shows simple correlations (r) between all regression variable pairs, with ratios of these to the corresponding r values for Southcenter. As shown, the ratios of r values for the two sites are quite variable and often of opposite sign for those variables which are modest to poor predictors (e.g.,

TABLE 36. RESULTS OF VALIDATION-GENERALIZATION (V-G) USING DATA FROM SOUTHCENTER AND COEFFICIENTS FROM VIEWRIDGE 1 (BOTH SEATTLE)*

Variable	Viewridge 1 R^2	V-G R^2	$\frac{\text{V-G } R^2}{\text{Viewridge 1 } R^2} \times 100\%$
Lead concentration	.792	.447	56.
Zinc concentration	.733	.545	74.
Nitrite plus nitrate N concentration	.754	.239	32.
Total phosphorus P	.678	.010	1.5
Organic nitrogen	.564	.461	81.7

*"Validation-generalization" is a term used for the application of a previously developed regression equation to a new set of data, as from a different site. The value of R^2 so obtained, when compared with R^2 from the new regression equation suggests the degree of consistency from data set to data set.

TABLE 37. REGRESSION EQUATIONS WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSIONS ONLY, α = .05, TWO-TAILED TEST) FROM SOUTHCENTER, SEATTLE, DATA

Equations	R^2
$[N_{org}] = (.00891 \overset{\times}{\div} 1.70)\ [Pb]^{(.735 \pm .103)}$	.503
$[N_{org}] = (.0109 \overset{\times}{\div} 1.82)\ [Zn]^{(.764 \pm .125)}$	.428
$[N_{org}] = (1.26 \overset{\times}{\div} 1.24)\ [NO_2 + NO_3]^{(.631 \pm .101)}$	.438
$\log_{10}\ [N_{org}] = 1.78 + .000680 + .477 \log_{10} [Pb] + .369 \log_{10} [NO_2 + NO_3]$	.626
$[NO_2 + NO_3] = (.00549 \overset{\times}{\div} 1.91)\ [Pb]^{(.654 \pm .122)}$	.363
$[NO_2 + NO_3] = (.00288 \overset{\times}{\div} 1.82)\ [Zn]^{(.858 \pm .124)}$	.491
$[NO_2 + NO_3] = (.305 \overset{\times}{\div} 1.17)\ [N_{org}]^{(.694 \pm .111)}$	.438
$[NO_2 + NO_3] = (2.40 \overset{\times}{\div} 1.66)\ [0\text{-}PO_4]^{(.787 \pm .143)}$	.377
$\log_{10}\ [NO_2 + NO_3] = -1.48 + .287 \log_{10} DD - .078\ HSS + .610 \log_{10} [Zn] + .338 \log_{10} [0\text{-}PO_4]$	.716
$[Pb] = (5.13 \overset{\times}{\div} 1.77)\ [Zn]^{(.750 \pm .118)}$	.444
$[Pb] = (479. \overset{\times}{\div} 1.26)\ [NO_2 + NO_3]^{(.554 \pm .104)}$	.362
$[Pb] = (324. \overset{\times}{\div} 1.15)\ [N_{org}]^{(.685 \pm .096)}$	.503
$[Pb] = (525. \overset{\times}{\div} 1.38)\ [P_{tot}]^{(.572 \pm .146)}$	.236
$\log_{10}\ [Pb] = 1.28 + .422 \log_{10} [SS] + .314 \log_{10} [Zn] + .247 \log_{10} [N_{org}] - .210\ CR$	.713
$[Zn] = (5.20 \overset{\times}{\div} 1.63)\ [Pb]^{(.591 \pm .093)}$	.444
$[Zn] = (309. \overset{\times}{\div} 1.17)\ [NO_2 + NO_3]^{(.572 \pm .082)}$	.491
$[Zn] = (182 \overset{\times}{\div} 1.12)\ [N_{org}]^{(.561 \pm .092)}$	.428
$[Zn] = (309. \overset{\times}{\div} 1.29)\ [P_{tot}]^{(.543 \pm .126)}$	.271
$\log_{10}\ [Zn] = 1.59 + 1.26D + .507 \log_{10} [NO_2 + NO_3] + .351 \log_{10} T$	.631
$\log_{10}\ [P_{tot}] = 2.79 - 1.30D + .201 \log_{10} [N_{org}] + .001\ C + .012\ T$	.612

TABLE 37 (continued)

Explanation of symbols:

N_{org} = organic nitrogen, ppm

Pb = lead, ppb

Zn = zinc, ppb

$NO_2 + NO_3$ = nitrite plus nitrate N, ppm

D = days since 1/1/73

$O\text{-}PO_4$ = orthophosphate phosphorus, ppm

DD = antecedent dry days (dry days preceding each storm)

HSS = hours since start of storm

P_{tot} = total phosphorus, ppm

SS = suspended solids, ppm

CR = cumulative rainfall per storm, inches

T = turbidity, JTU

C = conductivity, μmhos/cm

n = 199 for all equations

TABLE 38. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT AND INDEPENDENT VARIABLES (SOUTH SEATTLE)

	Days since 1/1/73*	Month*	Hours since storm start	Dry days before storm	Suspended solids	Lead concentration	Zinc concentration	Nitrite plus nitrate N concentration	Ammonia N concentration	Total phosphorus concentration	Orthophosphate phosphorus concentration	Organic nitrogen N concentration	Conductivity	Turbidity	Flow	Storm rainfall (cumulative)*
Lead	-.313 / -3.73	-.175 / -.440	-.131 / 0.46	.307 / 1.33	.819 / 1.23	1.00	.792 / 1.19	.542 / 0.90	.574 / 1.64	.744 / 1.53	.572 / 1.50	.515 / 0.73	.449 / 2.97	.769 / 1.40	.343 / -4.70	-.043 / 0.16
Zinc	-.221 / -1.09	-.332 / 8.30	-.146 / 0.72	.348 / 2.49	.625 / 1.67	.792 / 1.19	1.00	.636 / 0.91	.681 / 2.06	.708 / 1.36	.537 / 1.26	.593 / 0.91	.689 / 1.39	.529 / 1.08	.068 / -.026	-.274 / 0.88
Nitrite plus nitrate N	-.529 / 37.8	-.291 / -1.27	-.185 / 1.41	.538 / 1.11	.348 / 1.17	.542 / 0.90	.636 / .910	1.00	.751 / 1.29	.520 / 1.10	.594 / 0.97	.604 / 0.91	.642 / 1.53	.371 / 1.37	-.064 / 0.23	-.208 / 0.73
Ammonia N	-.384 / 1.53	-.366 / -.98	-.088 / -.83	.519 / 0.99	.415 / 3.07	.574 / 1.64	.681 / 2.06	.751 / 1.29	1.00	.586 / 2.83	.542 / 0.99	.551 / 1.93	.583 / 9.4	.446 / 5.95	-.029 / 0.23	-.198 / 0.94
Total Phosphorus	-.278 / 2.14	-.175 / 1.09	.085 / -.29	.336 / 2.43	.711 / 1.82	.744 / 1.53	.708 / 1.36	.520 / 1.10	.586 / 2.83	1.00	.657 / 1.22	.446 / 0.84	.539 / 1.01	.657 / 1.16	.240 / -0.88	-.040 / 0.17
Orthophosphate Phosphorus	-.175 / -1.55	.018 / .04	.142 / 2.03	.353 / 0.88	.493 / 4.40	.572 / 1.50	.537 / 1.26	.594 / 0.97	.542 / 0.99	.657 / 1.22	1.00	.449 / 0.93	.459 / 1.17	.529 / 1.53	.115 / -0.42	-.009 / 0.03
Organic Nitrogen	-.181 / -.080	-.249 / -.93	-.052 / 0.18	.258 / 0.97	.340 / 0.66	.515 / 0.73	.593 / 0.91	.604 / 0.91	.551 / 1.93	.446 / 0.84	.449 / 0.93	1.00	.548 / 1.73	.299 / 0.63	.016 / -0.10	-.184 / 0.77

Entries in table are $r_{\text{South Seattle}} \Big/ \frac{r_{\text{South Seattle}}}{r_{\text{Southcenter}}}$

All variables except those marked with "" are $\log_{10}$ transformed.

days since 1/1/73, month of the year, hours since start of storm, and flow). In contrast, good predictors (e.g., nitrite plus nitrate nitrogen, orthophosphate phosphorus, and organic nitrogen) give ratios fairly close to 1.0, suggesting similar degrees of correlation from site to site. Moderately good predictors (e.g., suspended solids and antecedent dry days) give positive r ratios (connoting relationships of consistent sign of slope) but ratios which are less nearly equal to 1.0. Thus, as might be expected, the degree of spatial consistency of correlations improves as r approaches ±1. That is, the better the predictor, the more spatially consistent the correlation between predictor and predicted variable.

Table 39 presents t statistics for multiple regression coefficients for South Seattle data. In four cases (except organic nitrogen), R^2 is higher in this data base than for Southcenter. The values are generally similar, however, with lead and zinc being among the best explained using the independent variables available. Also similar to Southcenter is the fact that every regression selected one or more nitrogen variables, and four selected one or more representing suspended matter. In this case, however, all five regressions used at least one heavy metal. In contrast, storm-related variables in this data base appeared to contribute relatively little, overall, to predicting runoff quality, whereas at Southcenter, at least one storm-related variable was selected for each regression.

Table 40 shows results of validation-generalization of South Seattle data using Viewridge 1 regression equations. Like Southcenter, the best validations are for heavy metals and organic nitrogen. Nitrate plus nitrite validates modestly, while total phosphorus validates poorly. This suggests that similar physical/chemical phenomena determine runoff lead, zinc, and organic nitrogen concentrations, and concentrations of nitrate plus nitrite, to an extent as well. In contrast, total phosphorus concentrations in runoff appear to be highly variable spatially both comparing Southcenter to Viewridge 1 and South Seattle to Viewridge 1 (based on validation R^2) and comparing Southcenter and South Seattle (based on variables used in the regressions).

Table 41 presents predictive equations developed for South Seattle data.

Seattle: Viewridge 1

Table 42 shows simple correlations (r) between dependent and independent variables for Viewridge 1 data, and ratios of r for Viewridge 1 to South Seattle and to Southcenter r values. The r values for correlations between water quality variables and the first three predictor variables ("days since 1/1/73," "month," and "hours since start of storm") are relatively low, as they are for Southcenter and South Seattle. Also, the ratios are quite variable suggesting not only weak, but also spatially inconsistent and possibly spurious relationships. Conductivity, "cumulative storm rainfall" and especially, flow, are similarly weak and spatially inconsistent predictors. In contrast, and again like South Seattle and Southcenter, "antecedent dry days," suspended solids, and turbidity are better runoff water quality predictors, with r being commonly in the 0.5-0.6 range. Here, as has been observed earlier, high r values are associated with spatially consistent values as shown by ratios fairly close to 1.0. Correlations among water quality variables, themselves, are

TABLE 39. MULTIPLE REGRESSION (t) STATISTICS AND R^2 FOR SOUTH SEATTLE DATA

	Lead concentration, µg/ℓ†	Zinc concentration, µg/ℓ†	Nitrite and nitrate concentration, mg/ℓ†	Organic nitrogen concentration, mg/ℓ	Total phosphorus concentration, mg/ℓ	
Days since 1/1/73			-9.29	2.76*		Long-term trends
Month of the year	*	*	*	*		Storm-related variables
Hours since start of storm	*				4.95	
Cumulative storm rainfall, inches	*	-3.02*	*	*	*	
Dry days before storm						
Suspended solids concentration, mg/ℓ	5.19	2.79				Suspended matter
Turbidity, JTU	6.02	-4.81	-5.16		4.85*	
Lead concentration, µg/ℓ		9.63		3.85	3.06	Heavy metals
Zinc concentration, µg/ℓ	9.33		3.69		3.36	
Nitrate and nitrite N concentration, mg/ℓ	2.48			5.22		Nitrogen variables
Ammonia N concentration, mg/ℓ		4.07	5.99		2.85	
Organic nitrogen concentration, mg/ℓ	2.59		4.00			
Total phosphorus concentration, mg/ℓ						Phosphorus variables
Orthophosphate P concentration, mg/ℓ			6.17			
Conductivity, µmhos/cm	-4.70	7.65	*	2.88	*	
Flow, cfs	*			*	*	
R^2	.850	.814	.762	4.67	.679	
n	199	199	199	199	199	
$n^{††}$	63	47	41	155	104	

*Variables not $\log_{10}$ transformed.

†Substantial autocorrelation of residuals was noted. Estimated number of independent observations very much less than n (199).

††Estimated n, based on a method due to Wold (See Appendix A) shown as $n^{††}$.

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at $\alpha = .05$, two-tailed test.

TABLE 40. RESULTS OF VALIDATION-GENERALIZATION (V-G) USING DATA FROM SOUTH SEATTLE AND COEFFICIENTS FROM VIEWRIDGE 1*

Variable	Viewridge 1 R^2	V-G R^2	$\frac{\text{V-G } R^2}{\text{Viewridge 1 } R^2} \times 100\%$
Lead concentration	.792	.688	86.9
Zinc concentration	.733	.605	82.5
Nitrate plus nitrite N concentration	.754	.322	42.7
Total phosphorus P	.678	.021	3.1
Organic nitrogen	.564	.333	59.0

*"Validation-generalization" is a term used for the application of a previously developed regression equation to a new set of data, as from a different site. The value of R^2 so obtained, when compared with R^2 from the new regression equation suggests the degree of consistency from data set to data set.

TABLE 41. REGRESSION EQUATIONS WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSIONS ONLY, $\alpha = .05$, TWO-TAILED TEST) FROM SOUTH SEATTLE DATA

Equations	R^2
$[N_{org}] = (.041 \overset{\times}{\div} 1.86)\ [Pb]^{(.526 \pm .124)}$	.266
$[N_{org}] = (.00490 \overset{\times}{\div} 2.5)\ [Zn]^{(.896 \pm .172)}$	.352
$[N_{org}] = (1.14 \overset{\times}{\div} 1.18)\ [NO_2 + NO_3]^{(.515 \pm .096)}$	.365
$[N_{org}] = (1.74 \overset{\times}{\div} 1.30)\ [NH_4]^{(.453 \pm .097)}$	.304
$[N_{org}] = (3.45 \overset{\times}{\div} 1.68)\ [OPO_4\text{-}P]^{(.562 \pm .158)}$	.202
$[N_{org}] = (.026 \overset{\times}{\div} 2.00)\ C^{(.764 \pm .165)}$	.301
$\log_{10} [N_{org}] = -1.52 + .0004\ D + .490 \log_{10} [Zn] + .403 \log_{10} [NO_2 + NO_3]$	.452
$[NO_2 + NO_3] = (3.3 \times 10^9 \overset{\times}{\div} 2.24 \times 10^2)\ D^{(-3.40 \pm .79)}$	.269
$[NO_2 + NO_3] = (.174 \overset{\times}{\div} .069)\ DD^{(.375 \pm .083)}$	.289
$[NO_2 + NO_3] = (.010 \overset{\times}{\div} 2.09)\ [Pb]^{(.649 \pm .142)}$	.294
$[NO_2 + NO_3] = (1.56 \overset{\times}{\div} 1.27)\ [NH_4]^{(.723 \pm .090)}$	.564
$[NO_2 + NO_3] = (6.61 \times 10^{-4} \overset{\times}{\div} 2.81)\ [Zn]^{(1.13 \pm .196)}$	.405
$[NO_2 + NO_3] = (.392 \overset{\times}{\div} 1.15)\ [N_{org}]^{(.708 \pm .132)}$	.365
$[NO_2 + NO_3] = (1.25 \overset{\times}{\div} 1.33)\ [P_{tot}]^{(.641 \pm .149)}$	.270
$[NO_2 + NO_3] = (4.26 \overset{\times}{\div} 1.73)\ [OPO_4\text{-}P]^{(.870 \pm .167)}$	.353
$\log_{10} [NO_2 + NO_3] = 6.12 - 1.95 \log_{10} D + .374 \log_{10} [NH_4] + .280 \log_{10} [N_{org}] + .328 \log_{10} [OPO_4\text{-}P]$	.723
$[Pb] = (.299 \overset{\times}{\div} 1.98)\ [Zn]^{(1.17 \pm .13)}$	.627
$[Pb] = (468. \overset{\times}{\div} 1.29)\ [NH_4]^{(.462 \pm .093)}$	.329
$[Pb] = (275. \overset{\times}{\div} 1.17)\ [NO_2 + NO_3]^{(.454 \pm .100)}$	.294
$[Pb] = (199. \overset{\times}{\div} 1.12)\ [N_{org}]^{(.505 \pm .119)}$	.266

TABLE 41 (continued)

Equations	R^2
$[Pb] = (575. \overset{\times}{\div} 1.23)\ [P_{tot}]^{(.766 \pm .097)}$	.553
$[Pb] = (1412. \overset{\times}{\div} 1.58)\ [OPO_4\text{-}P]^{(.701 \pm .142)}$	.327
$[Pb] = (12.9 \overset{\times}{\div} 2.05)\ C^{(.612 \pm .172)}$	.201
$[Pb] = (18.6 \overset{\times}{\div} 1.29)\ T^{(.752 \pm .088)}$	.592
$\log_{10} [Pb] = -.0047 + .163 \log_{10} [SS] + .823 \log_{10} [Zn] - .187 \log_{10} C + .301 \log_{10} T$	.835
$[Zn] = (13.8 \overset{\times}{\div} 1.35)\ [Pb]^{(.535 \pm .058)}$	.627
$[Zn] = (501. \overset{\times}{\div} 1.15)\ [NH_4]^{(.370 \pm .056)}$	.463
$[Zn] = (331. \overset{\times}{\div} 1.12)\ [NO_2 + NO_3]^{(.359 \pm .061)}$	.405
$[Zn] = (251. \overset{\times}{\div} 1.07)\ [N_{org}]^{(.393 \pm .076)}$	.352
$[Zn] = (479. \overset{\times}{\div} 1.15)\ [P_{tot}]^{(.492 \pm .069)}$	.501
$[Zn] = (832. \overset{\times}{\div} 1.38)\ [OPO_4\text{-}P]^{(.444 \pm .098)}$	.289
$[Zn] = (15.5 \overset{\times}{\div} 1.48)\ C^{(.635 \pm .095)}$	.474
$[Zn] = (77.6 \overset{\times}{\div} 1.29)\ T^{(.349 \pm .079)}$	.279
$\log_{10} [Zn] = .967 + .475 \log_{10} [Pb] + .091 \log_{10} [NH_4] + .327 \log_{10} C - .143 \log_{10} T$	.800
$[P_{tot}] = (.00468 \overset{\times}{\div} 1.58)\ [Pb]^{(.722 \pm .091)}$	.553
$[P_{tot}] = (7.94 \times 10^{-4} \overset{\times}{\div} 2.14)\ [Zn]^{(1.02 \pm .146)}$	.501
$[P_{tot}] = (.313 \overset{\times}{\div} 1.18)\ [NO_2 + NO_3]^{(.422 \pm .098)}$	.270
$[P_{tot}] = (.547 \overset{\times}{\div} 1.27)\ [NH_4]^{(.458 \pm .089)}$	.343
$\log_{10} [P_{tot}] = (-1.03 \pm .07) + (.0118 \pm .002)\ T$	.375
$\log_{10} [P_{tot}] = -2.65 + .00033\ HSS + .268 \log_{10} [Pb] + .498 \log_{10} [Zn] + .005\ T$	.666

Explanation of symbols on following page

TABLE 41 (continued)

Explanation of Symbols:

N_{org} = organic nitrogen, ppm

Pb = lead, ppb

Zn = zinc, ppb

$NO_2 + NO_3$ = nitrite + nitrate N, ppm

NH_4 = ammonia N, ppm

OPO_4-P = orthophosphate P, ppm

C = conductivity, µmhos/cm

D = days since 1/1/73

DD = antecedent dry days

P_{tot} = total phosphorus, ppm

T = turbidity, JTU

HSS = hours since start of storm

TABLE 42. SIMPLE CORRELATIONS (r) BETWEEN DEPENDENT AND INDEPENDENT VARIABLES (VIEWRIDGE 1)

	Days since 1/1/73	Month*	Hours since start of storm*	Dry days before storm	Suspended solids concentration	Lead concentration	Zinc concentration	Ammonia concentration	Nitrite + nitrate concentration	Organic nitrogen concentration	Total phosphorus concentration	Orthophosphate concentration	Conductivity*	Turbidity	Flow	Cumulative storm rainfall*
Lead concentration	-.333 1.06 -3.97	-.169 .968 -.910	-.355 2.71 1.25	.550 1.79 2.38	.655 .800 .984	1.00	.746 .942 1.12	.356 .619 1.02	-.216 -.399 -.359	.541 1.05 .766	.563 .757 1.16	.434 .759 1.14	-.253 -.563 -1.67	.614 .799 1.12	.137 .400 -1.88	-.051 1.19 .190
Zinc concentration	.073 -.332 .362	-.257 .775 6.44	-.367 2.52 1.81	.357 1.03 2.55	.537 .860 1.44	.746 .942 1.12	1.00	.315 .463 .953	-.016 -.026 -.023	.579 .976 .889	.480 .678 .921	.368 .685 .863	.013 .018 .026	.496 .938 1.01	-.034 -.499 .013	-.209 .762 .671
Ammonia concentration	-.284 .741 1.13	-.190 .519 -.508	-.059 .669 -.555	.530 1.021 1.011	.319 .769 2.36	.356 .619 1.02	.315 .463 .953	1.00	.438 .583 .753	.304 .552 1.07	.640 1.09 3.09	.685 1.26 1.25	.179 .468 4.40	.273 .612 3.64	-.110 3.78 .869	-.219 1.11 1.04
Nitrite + nitrate concentration	-.002 .005 .175	-.185 .637 -.808	.113 -.611 -.862	.089 .165 .183	-.214 -.615 -.719	-.216 -.399 -.359	-.016 -.026 -.023	.438 .583 .753	1.00	.176 .291 .265	.261 .502 .553	.345 .580 .563	.757 1.18 1.81	-.101 -.271 -.371	-.525 8.21 1.89	-.169 .814 .594
Organic nitrogen concentration	-.143 .791 -.063	-.321 1.29 -1.20	-.248 4.78 .860	.237 .919 .891	.513 1.51 .996	.541 1.05 .766	.579 .976 .889	.304 .552 1.07	.176 .291 .265	1.00	.487 1.09 .98	.326 .725 .675	.234 .427 .738	.528 1.77 1.11	-.127 -7.94 .794	-.075 .407 .313
Total phosphorus concentration	-.337 1.21 2.60	-.344 1.97 2.14	-.288 -3.39 .982	.576 1.72 4.17	.419 .589 1.07	.563 .757 1.16	.480 .678 .921	.640 1.09 3.09	.261 .502 .553	.487 1.09 .918	1.00	.715 1.09 1.33	.189 .350 .353	.454 .691 .802	-.129 -.539 .474	-.342 8.55 1.45
Orthophosphate phosphorus concentration	.291 -1.66 2.58	-.198 -11.0 -.439	.154 1.09 2.21	-.048 -.136 -.120	-.255 -.518 -2.28	.434 .759 1.14	.368 .685 .863	.685 1.26 1.25	.345 .580 .563	.326 .725 .675	.715 1.09 1.33	1.00	.105 .230 .269	.299 .565 .864	-.131 -1.14 .478	-.035 3.89 .117

Entries in table are r for Viewridge 1, ratio of r for Viewridge 1 to r for South Seattle (above diagonal), and ratio of r for Viewridge 1 to r for Southcenter (below diagonal).

All variables except those marked with '' are $\log_{10}$ transformed.

again high (although not as high overall as at South Seattle), with many values in the 0.5-0.7 range, and with r ratios comparing Viewridge 1 correlations with those of Southcenter and South Seattle often being near 1.0. Notably, whereas nitrite plus nitrate was a good predictor for other water quality constituents at Southcenter and South Seattle, and r ratios were close to 1.0, this variable is a poor predictor of other water quality variables at Viewridge 1.

Table 43 presents multiple regression statistics for Viewridge 1. Like Southcenter and South Seattle, it appears that for a given regression, variables are selected from among the general groups (e.g., heavy metals, nitrogen variables), with only one case of more than one from any group (i.e., suspended solids and turbidity as predictors of organic nitrogen).

In all five regressions, a heavy metal variable was selected (lead, zinc, or cadmium) and four out of five selected a suspended solids variable. In contrast to South Seattle and Southcenter, nitrogen variables seem relatively less important (as evidenced by only three regressions using a nitrogen variable), and the occurrence of "days since 1/1/73" in all five regressions suggests that conditions at Viewridge 1 are changing over time.

Regression R^2 values for Viewridge 1 were high (all greater than 0.7 except for organic nitrogen) as they were at South Seattle and Southcenter. At all three sites, regression R^2 values for lead, zinc, and nitrite plus nitrate were in excess of 0.7. Generally, the value of R^2 was greater for total phosphorus than for organic nitrogen which was lowest (except at Southcenter).

Table 44 presents regression equations for Viewridge 1 data.

TABLE 43. MULTIPLE REGRESSION (t) STATISTICS AND R^2 FOR VIEWRIDGE 1 (SEATTLE) DATA

	Lead concentration, μg/ℓ	Zinc concentration, μg/ℓ[†]	Nitrite and nitrate N concentration, mg/ℓ	Organic nitrogen concentration, mg/ℓ[†]	Total phosphorus concentration, mg/ℓ	
Days since 1/1/73	-5.30	9.22	-2.46	-3.24	-4.18	Long-term trends
Month of the year	*	*	*	*	*	Storm-related variables
Hours since start of storm	*	*	*	*	*	
Cumulative storm rainfall, inches	*	*	*	*	-7.73*	
Dry days before storm	2.70			-2.50		
Suspended solids concentration, mg/ℓ	3.01		-3.63	4.06	2.27	Suspended matter
Turbidity, JTU				3.20		
Lead concentration, μg/ℓ		14.3	-4.65**		4.89	Heavy metals
Zinc concentration, μg/ℓ	12.1	.		5.65		
Nitrate and nitrite N concentration, mg/ℓ	-5.95					Nitrogen variables
Ammonia N concentration, mg/ℓ			7.83			
Organic nitrogen concentration, mg/ℓ		3.91				
Total phosphorus concentration, mg/ℓ	2.84					Phosphorus variables
Orthophosphate P concentration, mg/ℓ		2.56			10.0	
Conductivity, μmhos/cm	*	*	12.3 *	7.24*	5.33*	
Flow, cfs		-2.44	-3.01			
R^2	.792	.733	.753	.564	.741	
n	188	188	188	188	188	
n[††]	188	30	188	46	150	

*Variables not $\log_{10}$ transformed.

**Cadmium data which were not used in other regressions, correlated with this variable, and this is the value of t for the cadmium coefficient (μg/ℓ).

[†] Substantial autocorrelation of residuals was noted. Estimated number of independent observations very much less than n (188).

[††] Estimated n shown as n[††]- based upon as estimation due to Wold (see Appendix A).

Table entries are values of t for the coefficients. Selection of regressions was made such that the greatest number of predictors was used subject to at least 30 observations per predictor variable and all t significant at α = .05, two-tailed test.

TABLE 44. REGRESSION EQUATIONS, WITH CONFIDENCE INTERVALS FOR SLOPE AND INTERCEPT (SIMPLE REGRESSIONS ONLY, $\alpha = .05$, TWO-TAILED TEST) FROM VIEWRIDGE 1 (SEATTLE) DATA

Equations	R^2
$[N_{org}] = (.130 \stackrel{\times}{\div} 1.56)\ [Pb]^{(.394 \pm .089)}$	.292
$[N_{org}] = (.066 \stackrel{\times}{\div} 1.70)\ [Zn]^{(.600 \pm .123)}$	.335
$[N_{org}] = (1.63 \stackrel{\times}{\div} 1.21)\ [P_{tot}]^{(.470 \pm .123)}$	.238
$[N_{org}] = (.288 \stackrel{\times}{\div} 1.31)\ T^{(.463 \pm .108)}$	.279
$\log_{10} [N_{org}] = -1.29 + .161 \log_{10} SS + .294 \log_{10} [Zn] + .002\ C + .234 \log_{10} T$	.537
$\log_{10} [NO_2 + NO_3] = (-.869 \pm .079) + (.0051 \pm .0006)\ C$	.574
$[NO_2 + NO_3] = (.888 \stackrel{\times}{\div} 1.19)\ Q^{(-.538 \pm .127)}$	.276
$[Pb] = (83.2 \stackrel{\times}{\div} 1.15)\ DD^{(.376 \pm .083)}$	.302
$[Pb] = (1.30 \stackrel{\times}{\div} 1.07)\ [Zn]^{(1.06 \pm .137)}$	.557
$[Pb] = (135. \stackrel{\times}{\div} 1.12)\ [N_{org}]^{(.741 \pm .167)}$	.292
$[Pb] = (339. \stackrel{\times}{\div} 1.29)\ [P_{tot}]^{(.744 \pm .159)}$	.317
$[Pb] = (21.4 \stackrel{\times}{\div} 1.41)\ T^{(.739 \pm .138)}$	.377
$\log_{10} [Pb] = 6.05 - 1.90\ D + .947 \log_{10} [Zn] -.296 \log_{10} [NO_2 + NO_3] + .272 \log_{10} [P_{tot}]$	.774
$[Zn] = (5.78 \stackrel{\times}{\div} 1.40)\ [Pb]^{(.525 \pm .068)}$	.557
$[Zn] = (77.6 \stackrel{\times}{\div} 1.07)\ [N_{org}]^{(.559 \pm .115)}$	.335
$[Zn] = (132. \stackrel{\times}{\div} 1.20)\ [P_{tot}]^{(.446 \pm .119)}$	.230
$[Zn] = (26.9 \stackrel{\times}{\div} 1.32)\ T^{(.420 \pm .107)}$	.246
$\log_{10} [Zn] = 3.93 + 1.65 \log_{10} D + .494 \log_{10} [Pb] + .203 \log_{10} [N_{org}] + .120 \log_{10} [OPO_4\text{-}P]$	.725

TABLE 44 (continued)

Explanation of Symbols:

N_{org} = organic nitrogen, ppm

Pb = lead, ppb

Zn = zinc, ppb

P_{tot} = total phosphorus

T = turbidity, JTU

SS = suspended solids, ppm

C = conductivity, µmhos/cm

$NO_2 + NO_3$ = nitrite plus nitrate N, ppm

Q = flow, cfs

DD = antecedent dry days

D = days since 1/1/73

OPO_4-P = orthophosphate phosphorus

REFERENCES

1. Metcalf and Eddy, Inc., University of Florida, and Water Resources Engineers, Inc. Storm Water Management Model (four volumes) 11024 DOC 07/71, 11024 DOC 08/71, 11024 DOC 09/71, and 11024 DOC 10/71, U.S. Environmental Protection Agency, Water Quality Office, 1971.

2. Donigian, A.S., Jr., and N.H. Crawford. Modeling Nonpoint Pollution from the Land Surface. EPA-600/3-76-083, U.S. Environmental Protection Agency, Athens, Georgia. July 1976.

3. U.S. Army Corps of Engineers. Storage, Treatment, Overflow Runoff Model "STORM." 723-S8-L7520, Hydrologic Engineering Center, Davis, California. 1976.

4. Donigian, A.S., Jr., and N.H. Crawford. Modeling Pesticides and Nutrients on Agricultural Lands. EPA-600/2-76-043, U.S. Environmental Protection Agency, Athens, Georgia. February 1976.

5. Hydrocomp. Model currently under development under the support of U.S. Environmental Protection Agency, Athens, Georgia. Work performed at Hydrocomp, Inc., Palo Alto, California.

6. Donigian, A.S., Jr., and N.H. Crawford. Simulation of Nutrient Loadings in Surface Runoff with the NPS Model. EPA-600/3-77-065, U.S. Environmental Protection Agency, Athens, Georgia. June 1977.

7. Cowen, W.F., K. Sirisinha, and G.F. Lee. "Nitrogen Availability in Urban Runoff." Journal WPCF, Vol. 48, no. 2, February 1976. Pp. 339-45.

8. Cowen, W.F. and G.F. Lee. "Phosphorus Availability in Particulate Materials Transported by Urban Runoff." Journal WPCF, Vol. 48, no. 3, March 1976. Pp. 580-91.

9. Sartor, J.D., G.B. Boyd, and F.J. Agardy. "Water Pollution Aspects of Street Surface Contaminants." Journal WPCF, Vol. 46, no. 3, March 1974. Pp. 458-67.

10. Colston, N.V., Jr. Characterization and Treatment of Urban Land Runoff. EPA-670/2-74-096, U.S. Environmental Protection Agency, Cincinnati, Ohio. December 1974.

11. Smith, C.N., R.A. Leonard, G.W. Langdale, and G.W. Bailey. *Transport of Agricultural Chemicals from Small Upland Piedmont Watersheds*. EPA-600/3-78-056, U.S. Environmental Protection Agency, Athens, Georgia, and U.S. Department of Agriculture, Watkinsville, Georgia. May 1978.

12. Ellis, B.G., A.E. Erickson, A.R. Wolcott, M. Zabik, and R. Leavitt. *Pesticide Runoff Losses from Small Watersheds in Great Lakes Basin.* EPA-600/3-77-112, U.S. Environmental Protection Agency, Athens, Georgia. October 1977.

13. Römkens, M.J.M., D.W. Nelson, and J.V. Mannering. "Nitrogen and Phosphorus Composition of Surface Runoff as Affected by Tillage Method." J. Environ. Quality, Vol. 2, no. 2, 1973. Pp. 292-5.

14. Burwell, R.E., G.E. Schuman, K.E. Saxton, and H.G. Heinemann. "Nitrogen in Subsurface Discharge from Agricultural Watersheds." J. Environ. Quality, Vol. 5, no. 3, 1976. Pp. 325-9.

15. Burwell, R.E., D.R. Timmons, and R.F. Holt. "Nutrient Transportation in Surface Runoff as Influenced by Soil Cover and Seasonal Periods." Soil Sci. Soc. Amer. Proc., Vol. 39, 1975. Pp. 523-8.

16. Morris, R.L. and L. Johnson. *Buffalo Bill Watershed Agricultural Run-Off Study*. Preliminary Report, Iowa State Hygienic Laboratory (University of Iowa), Iowa City, Iowa. February 1974.

17. Ellis, B.G., A.E. Erickson, and A.R. Wolcott. *Nitrate and Phosphorus Runoff Losses from Small Watersheds in Great Lakes Basin*. EPA-600/3-78-028, U.S. Environmental Protection Agency, Athens, Georgia. March 1978.

18. Iwatsubo, R.T., K.M. Nolan, D.R. Harden, G.D. Glysson, and R.J. Janda. *Redwood National Park Studies, Data Release Number 1, Redwood Creek, Humboldt County, California, September 1, 1973 - April 10, 1974*. U.S. Geological Survey, Menlo Park, California. December 1975.

19. Iwatsubo, R.T., K.M. Nolan, D.R. Harden, and G.D. Glysson, *Redwood National Park Studies, Data Release Number 2, Redwood Creek, Humboldt County, and Mill Creek, Del Norte County, California, April 11, 1974 - September 30, 1975*. U.S. Geological Survey, Menlo Park, California. December 1976.

20. Huber, W.C. and J.P. Heaney. *Urban Rainfall-Runoff-Quality Data Base.* EPA-600/8-77-009, U.S. Environmental Protection Agency, Cincinnati, Ohio. July 1977.

21. Heidelburg College River Studies Laboratory. *Water Quality Data for Sandusky River Material Transport Stations*. U.S. Army Corps of Engineers, Buffalo, New York. Undated.

APPENDIX A

REGRESSION ANALYSIS THEORY

In order to understand the results of analysis presented in the body of this report, it is necessary to understand the statistical methods used, what the individual specific results suggest about water quality relationships, what the statistical assumptions mean, and what the techniques employed can and cannot do.

A considerable risk is associated with the interpretation of statistical results if the implications of assumptions and details of techniques are not understood. For example, it is always possible to fit a straight line to data consisting of three or more coplanar points. The slope and intercept will seldom be exactly zero, and as a result, it is all too easy to assume that the fitted line has some meaning when it, in fact, does not. It is through proper interpretation of statistical results that significant relationships are distinguished from the non-significant relationships.

In general, regression analysis may be considered as a two part process. The first part is nonstatistical in the sense that probabilities are not involved and assumptions about data distributions are not made. This is the fitting of a model (a predictive equation) to the data. The second part involves making assumptions about the underlying distributions from which the data are drawn. Where the assumptions are valid, statements can be made about the probability that relationships are real or that apparent relationships have probably occurred by chance alone.

MODEL FITTING

The first step in regression analysis consists of deciding upon the model to be used based upon assumed or known relationships in the data. Consider, first, a simple linear model which is assumed to represent some data set.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \text{(A-1)}$$

where Y_i = ith observed value of the dependent variable

X_i = ith observed value of the independent variable

β_0, β_1 = coefficients representing the true population relationship

ε_i = ith error term

The equation to be fitted to the data is:

$$\hat{Y}_i = b_0 + b_1 X_i \qquad \text{(A-2)}$$

where $\hat{Y}_i$ = ith predicted value for the dependent variable

b_0, b_1 = fitted coefficients and estimates of β_0 and β_1, respectively

Now the "least-squares prescription" specifies that the total sum of squares of deviations of the data values from the fitted line be minimized. To do this, first express the sum of squared deviations (or errors):

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2 = \sum_{i=1}^{n} (b_0 + b_1 X_i - Y_i)^2 = f(x,y) \qquad \text{(A-3)}$$

where n = the number of observations

Set the partials of the function with respect to b_0 and b_1 equal to zero:

$$\frac{\partial f(x,y)}{\partial b_0} = 2 \sum_{i=1}^{n} b_0 + 2b_1 \sum_{i=1}^{n} X_i - 2 \sum_{i=1}^{n} Y_i = 0 \qquad \text{(A-4)}$$

or

$$nb_0 + b_1 \sum_{i=1}^{n} X_i = \sum_{i-1}^{n} Y_i \qquad \text{(A-5)}$$

and

$$\frac{\partial f(x,y)}{\partial b_1} = 2b_0 \sum_{i=1}^{n} X_i + 2b_1 \sum_{i=1}^{n} X_i^2 - 2 \sum_{i=1}^{n} X_i Y_i = 0 \qquad \text{(A-6)}$$

or

$$b_0 \sum_{i=1}^{n} X_i + b_1 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i \qquad \text{(A-7)}$$

Equations (A-5) and (A-7) are called "normal equations," and solved simultaneously to give the coefficients b_0 and b_1.

The extension to multiple regression is straightforward. For any number of independent variables with a linear model, the equation to be fitted is:

$$Y = b_0 + b_1 X_{1,i} + b_2 X_{2,i} + \cdots + b_m X_{m,i} \qquad \text{(A-8)}$$

Solving for the matrix of b coefficient values is commonly done in a somewhat different way. In one approach, the matrix of correlation coefficients (r) is found as follows.

$$\begin{bmatrix} r_{1\cdot 1} & r_{1\cdot 2} & \cdots & r_{1\cdot m} \\ r_{2\cdot 1} & r_{2\cdot 2} & \cdots & r_{2\cdot m} \\ \vdots & \vdots & & \vdots \\ r_{m\cdot 1} & r_{m\cdot 2} & \cdots & r_{m\cdot m} \end{bmatrix} = \underline{r}_{i\cdot j} \qquad \text{(A-9)}$$

where if $r_{i\cdot j}$ is an element of $\underline{r}_{m\cdot m}$, then $r_{i\cdot j}$ is given by:

$$r_{i\cdot j} = \frac{\Sigma x_k x_\ell}{\sqrt{\Sigma x_k^2 \cdot \Sigma x_\ell^2}}, \quad \Sigma x_k x_\ell = \sum_{h=1}^{n} X_{k,h} X_{\ell,h} - \frac{\sum_{h=1}^{n} X_{k,h} \sum_{h=1}^{n} X_{\ell,h}}{n}, \quad \Sigma x_k^2 = \sum_{h=1}^{n} X_{k,h}^{\;2} - \frac{\left(\sum_{h=1}^{n} X_{k,h}\right)^2}{n} \qquad \text{(A-10)}$$

and for i=j, $r_{i\cdot j} = 1$.

Then the system of normal equations for the determination of the b's is:

$$\underline{r}_{m\cdot m}\ \underline{\beta}_m = \underline{r}_{y\cdot m} \qquad \text{(A-11)}$$

or

$$\underline{\beta}_m = \underline{r}_{y\cdot m}\ \underline{r}_{m\cdot m}^{-1} \qquad \text{(A-12)}$$

where $\underline{r}_{m\cdot m}$ = the symmetric matrix of correlation coefficients among the independent variables as defined by equations (A-9) and (A-10)

$\underline{\beta}_m$ = the column matrix of b to be found for equation (A-8)

$\underline{r}_{y\cdot m}$ = the column matrix of correlation coefficients between the dependent variable, y and each of the independent variables

Computation of the constant term, b_0 is performed as follows:

$$b_0 = \overline{Y} - b_1\overline{X}_1 - b_2\overline{X}_2 - \cdots - b_m\overline{X}_m \qquad \text{(A-13)}$$

It is clear that conceptually, the procedure for fitting any model by least squares is equally straightforward and consists of the following steps.

1. Define the model relating the dependent variable to the independent variable(s).
2. Define the function to be minimized (sum of squared deviations of observed values from the fitted line, surface, etc.).
3. Take the partials with respect to each model parameter and set equal to zero.
4. Solve the resulting equations simultaneously to give values for model coefficients.

However, depending upon the form of the equation, it may be difficult to explicitly solve for the desired coefficients.

It is customary as a next step, in any regression other than simple linear regression, to compute R^2, the squared multiple correlation coefficient. In simple linear regression (fitting an equation of the form $Y = b_0 + b_1X$), $R^2 \equiv r^2$. R^2 can be interpreted simply as the proportion of the total variance of the dependent variable which may be accounted for by the model. This is shown schematically in Figure A-1. R^2 may be computed as follows:

$$R^2 = \frac{SS_{Reg}}{SS_T} \tag{A-14}$$

where SS_{Reg} = sum of squares due to regression

SS_T = total sum of squares

$$SS = \sum_{i=1}^{n} Y_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} \tag{A-15}$$

where Y_i = predicted (for "Reg") and observed (for "T") values of the dependent variable

It should be noted that to this point nothing has been defined in terms of probabilities or "statistics." The fitting of a model by least squares is a purely mathematical process, and embodies no assumptions about underlying distributions. Similarly, the computation of R^2 is entirely devoid of assumptions about distributions and probabilities.

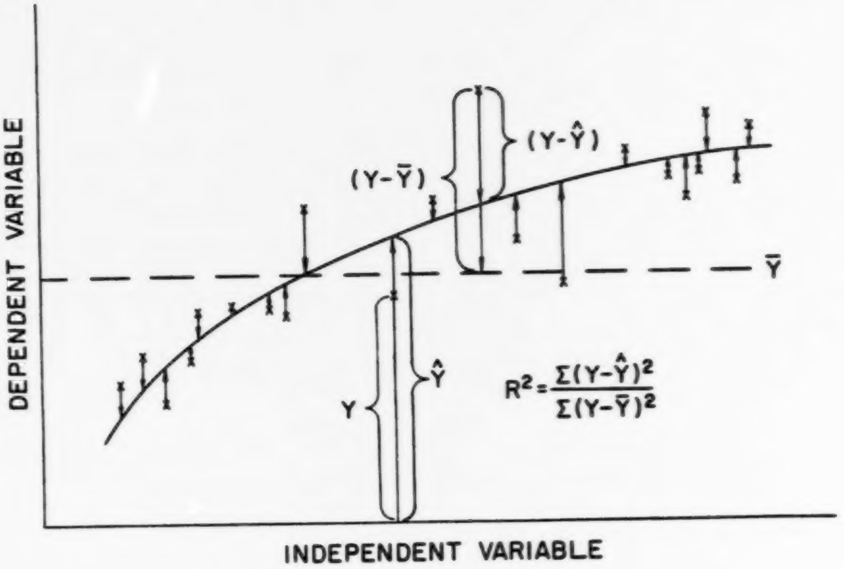

DEPENDENT VARIABLE
INDEPENDENT VARIABLE
(Y-Ŷ)
(Y-Ȳ)
Ȳ
Y
Ŷ
$R^2 = \frac{\Sigma(Y-\hat{Y})^2}{\Sigma(Y-\bar{Y})^2}$

Figure A-1. Schematic showing an interpretation of R^2.

REGRESSION STATISTICS

A number of important statistics are ordinarily generated in regression analyses. These include:

- F-statistic for the overall model.
- F-statistic for improvement in R^2 for each step in stepwise multiple regression.
- t-statistic for the coefficients or b-weights.
- Confidence intervals about the simple linear regression line for predicting individual values of the dependent variable or the mean, for a given value of the independent variable.
- Confidence intervals for the slope and intercept in simple linear regression.
- Partial correlation coefficients.
- Autocorrelation coefficients as a measure of independence of residuals.
- Confidence intervals about predicted values of the dependent variable in multiple regression.

Tests of hypotheses (the fundamental utility of computed statistics) involve assumptions about underlying distributions and characteristics of the data. The assumptions are:

- The data, for a given value of the independent variable, are normally distributed.
- The error component is independently distributed for a given value of the independent variable.
- The error variances are constant (homoscedastic) for different values of the independent variable.
- The independent variable has no error associated with it.

The last assumption is specific for certain regression applications. The first three assumptions may be summarized as follows: Given that:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i, \ i=1, n \tag{A-16}$$

where n = the number of observations.

Y_i = the value of the dependent variable for each value of X_i

ε_i = The error component of Y_i after accounting for the deterministic component, $b_0 + b_1 X_i$.

then the assumptions specify that the ε_i are normally and independently distributed with mean equal to zero and a standard deviation of σ_ε. $\sigma_\varepsilon{}^2$ is, in turn, uniform over X.

The F-Statistic

Provided the requisite assumptions are met, a number of important statistics may be generated and interpreted. Probably the most important of these is the F-statistic. The F-statistic for the overall regression is computed as

$$F = \frac{R^2 / df_N}{(1-R_2) / df_D} = \frac{SS_r / k}{SS_e / (n-k-1)} \qquad \text{(A-17)}$$

where df_N = degrees of freedom for the numerator = the number of independent variables (k) in the regression

df_D = degrees of freedom for the denominator = the number of observations (n) minus the number of independent variables (k) minus one

$(1-R^2)$ = the coefficient of non-determination

$$\left.\begin{array}{l} SS_r = \text{regression sum of squares} \\ SS_e = \text{total sum of squares} \end{array}\right\} \text{See equation (A-15)}$$

k = number of independent (predictor) variables

n = number of observations

The value of F is compared with $F_{critical}$ from a table of the F distribution. If $F > F_{critical}$, then the regression relationship is a significantly better estimate than is the mean at the specific confidence level.

F for comparing any step in the regression with the prior step is given by:

$$F = \frac{\left(R^2_{y\cdot 1,2,\ \cdot\ \cdot\ \cdot\ ,k} - R^2_{y\cdot 1,2,\ \cdot\ \cdot\ \cdot\ ,k-1}\right) / Df_N}{\left(1-R^2_{y\cdot 1,2,\ \cdot\ \cdot\ \cdot\ k}\right) / Df_D} \qquad \text{(A-18)}$$

where $R^2_{y\cdot 1,2,\ \cdot\ \cdot\ \cdot\ ,k}$ = R^2 for the current regression step

$R^2_{y\cdot 1,2,\ \cdot\ \cdot\ \cdot\ ,k-1}$ = R^2 for the preceding step

Df_N = 1

Df_D = the number of observations (n) minus the number of independent variables (k) minus one

The t-Statistic

With respect to regression, the t-statistic is commonly used for determining whether each coefficient in multiple regression is significant. If so, then at the selected confidence level, there is reason to believe the value is truly non-zero. If it is non-zero, then the best estimate of the coefficient is that determined in the regression. If the coefficient is not significant, then there is no reason to believe it is non-zero, and the variable should not be entered into the regression.

The t-statistic for each coefficient is given by:

$$t_j = b_j / SE_{bj} \qquad \text{(A-19)}$$

where b_j = the coefficient

SE_{bj} = the standard error of the coefficient

$$SE_{bj} = \sqrt{\frac{SE^2_{est}}{SS_{xj}\,(1-R^2_j)}} \qquad \text{(A-20)}$$

where SS_{xj} = the sum of squares of variable j

R^2_j = the squared multiple correlation coefficient regressing the jth variable on the remaining independent variables

$$SE_{est} = \sqrt{\frac{SS_{res}}{n-k-1}} \qquad \text{(A-21)}$$

where SE_{est} = standard error of estimate

SS_{res} = the residual sum of squares (sum of squares of differences between predicted Y and observed Y)

To estimate the confidence limits for predicting individual Y values and $\overline{Y}$ for a given value of X, where u and d designate upper and lower limits,

$$Y_{u,d} = Y \pm tS_y \qquad \text{(A-22)}$$

where S_y = standard error of sample mean, $\overline{Y}$ for a given value of X:

$$S_y = \sqrt{S_E^2 \left[\frac{1}{n} + \frac{(X-\bar{X})^2}{\Sigma x^2}\right]} \qquad \text{(A-23)}$$

or for S_y = standard error of individual Y values for a given X value:

$$S_y = \sqrt{S_E^2 \left[1 + \frac{1}{n} + \frac{(X-\bar{X})^2}{\Sigma x^2}\right]} \qquad \text{(A-24)}$$

In equations A-23 and A-24, Σx^2 is the sum of squares of deviations of X:

$$\Sigma x^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{n} \qquad \text{(A-25)}$$

Using these equations (A-22 through A-25) it is possible to establish confidence limits about the fitted line.

Partial Correlations

The partial correlation coefficients are useful in that they permit correlations of variables to be examined while removing the effect of any number of other variables on that correlation. For example, if it is desired to examine the correlation between variable 1 and variable 2, but it is suspected that some of the correlation of variables 1 and 2 is contaminated with variable 3, the effect of variable 3 may be removed from variables 1 and 2.

An example will clarify the utility of the partial correlations. Suppose we are interested in the correlation between the amount of a pesticide remaining on soil as a function of time with losses due to effects other than washoff. Correlating pesticide-remaining with time also includes the effect of stormwater runoff transport. Partialling out cumulative rainfall (or other rainfall variable) from time gives the correlation between time (and not rainfall) and the amount of pesticide remaining on the soil. If the correlation is substantial, this suggests a real loss of pesticide on soil over time due to factors other than washoff, such as biodegradation, photolysis, erosion, and volatilization.

The computation of the partial correlation coefficient is performed as follows:

$$r_{ij \cdot k} = \frac{r_{ij} - r_{ik} r_{jk}}{\sqrt{1 - r_{ik}^2}\,\sqrt{1 - r_{jk}^2}} \qquad \text{(A-26)}$$

where $r_{ij \cdot k}$ = the partial correlation coefficient for the correlation between variables i and j, both independent of variable k

Autocorrelation Coefficients

Autocorrelation of residuals and its evaluation are very important in regression analysis, particularly where time-series are involved. Autocorrelation statistics measure the degree of serial correlation in a data series. Thus it can be considered as $r_{x_i x_{i+j}}$ where if x represents individual observations, n is the number of such observations, and j is the lag, then i varies from 1 to n-j.

The autocorrelation coefficient, which can be designated as r_a is given by

$$r_a = \frac{\Sigma d_t d_{t+j}}{\Sigma d_t^2} \tag{A-27}$$

where d = residual (difference between observed and predicted value of Y)

j = selected lag

Clearly j can be set to 1 for the smallest degree of lag in the observations or it may be increased. If there is a periodicity to the autocorrelation, then r_a is likely to be similarly periodic.

Another measure of autocorrelation is the von Neumann Ratio:

$$\text{V.N.R.} = \frac{\dfrac{\Sigma \left(d_{t+j} - d_t\right)^2}{n - j}}{\dfrac{\Sigma d_t^2}{n}} \tag{A-28}$$

In general, assessments of autocorrelation using the two statistics tend to be consistent. In both cases, the statistic is computed and the value compared with tabulated critical values. Within a defined region, significant autocorrelation is indicated.

In regression, observations are assumed to be independent as discussed earlier, and evaluation of statistical results is predicated upon some numbers of degrees of freedom. Degrees of freedom are based, in turn, upon numbers of observations and variables. If observations are not independent, then the number of degrees of freedom, or, in essence, the number of sets of truly independent measures, can be severely overestimated, and data relationships may be judged meaningful when they are not.

Several methods have been cited in the literature for dealing with the problem of autocorrelation of residuals in time-series. Ezekiel and Fox*, for example have cited the use of "first differences" between observations rather than the observations themselves. This technique, which has been used successfully with econometric data, usually eliminates the autocorrelation of time-series regression residuals. However, in the author's experience, the technique can also seriously affect results of correlation analysis, and the value of the method is sensitive to features of the data. An example will clarify this. Suppose the following is a data set of observations and first differences with results of regression (Y = a + bX) as shown.

i	X_i	Y_i	$\hat{Y}_i$	d_i	d_i^2	$d_i\, d_{i+1}$	ΔX_i	ΔY_i	$\Delta\hat{Y}_i$	d_i	d_i^2	$d_i\, d_{i+1}$
1	1.00	37.3	39.03	1.73	2.99	-.208	0.02	1.9	0.98	-.92	.846	-2.37
2	1.02	39.2	39.08	-0.12	0.014	-.18	0.01	-1.6	0.98	2.58	6.66	-2.61
3	1.03	37.6	39.10	1.50	2.25	-.66	0.03	2.0	0.99	-1.01	1.02	-.48
4	1.06	39.6	39.16	-0.44	.194	.392	0.02	0.5	0.98	0.48	.23	.46
5	1.08	40.1	39.21	-0.89	.792	-2.74	2.13	0.7	1.65	0.95	.90	.27
6	3.21	40.8	43.88	3.08	9.49	7.45	0.02	0.7	0.98	0.28	.08	-1.26
7	3.23	41.5	43.92	2.42	5.86	-7.28	0.03	5.5	0.99	-4.51	20.3	-8.07
8	3.26	47.0	43.99	-3.01	9.06	6.32	0.05	-0.8	0.99	1.79	3.20	-1.07
9	3.31	46.2	44.10	-2.10	4.41	7.37	0.09	1.6	1.00	-0.60	.36	.01
10	3.40	47.8	44.29	-3.51	12.32	-9.20	3.80	2.2	2.18	-0.02	0.0	.08
11	7.20	50.0	52.62	2.62	6.86	-2.78	0.60	5.0	1.17	-3.83	14.7	-3.10
12	7.80	55.0	53.94	-1.06	1.12	1.10	0.10	0.2	1.01	0.81	.66	1.06
13	7.90	55.2	54.16	-1.04	1.082	.541	0.10	-0.3	1.01	1.31	1.72	3.55
14	8.00	54.9	54.38	-0.52	.270	-.728	0.10	-1.7	1.01	2.71	7.34	--
15	8.10	53.2	54.60	1.40	1.96	--	--	--	--	--	--	--
	a = 36.84			r_a = -0.0102			a = .9756		r_a = -.233			
	b = 2.192			(r_a not significant			b = .3157		(r_a not significant			
	r^2 = .9070			at 95% level)			r^2 = .0263		at 95% level)			
	F = 126.8						F = .325					

Regardless of whether the residuals are autocorrelated (they are not, in this case), it is clear that taking differences can cause a severe reduction in the correlation (r) in the data. This is particularly true when observations tend to be clustered in a number of groups, as in data taken frequently over time but only during storm activity. Because of this, the technique of using first differences, as is sometimes suggested, may not be useful in many analyses.

---

*Ezekiel, M. and K.A. Fox. 1959. Methods of Correlation and Regression Analysis. John Wiley and Sons, New York. pp. 340 et seq.

Another alternative is that suggested by Wold* and reported by Ezekiel and Fox (cited earlier in this Appendix). In this approach, the degrees of freedom are corrected using the formula

$$N' = N \Big/ \sqrt{1 + 2r_{a_1} + 2r_{a_2} + \cdot \cdot \cdot \; 2r_{a_n}} \qquad \text{(A-29)}$$

where N' = number of independent observations

N = number of observations

r_{a_j} = the autocorrelation coefficient for lag j.

This is the technique used in this study.

Confidence Intervals in Multiple Regression

In simple regression, confidence limits are constructed on Y for each value of X, permitting confidence envelopes to be described about the fitted line. In a similar fashion, it is possible to generate confidence interval estimates of predictions for each set of values for the independent variables in multiple regression. Although the process is conceptually the same, visualizing it becomes difficult because of the multidimensional nature of multiple regression. In three dimensions, the confidence limits consist of two curved surfaces located above and below the fitted surface. The formulas for the computation are presented in Ezekiel and Fox (cited earlier in this Appendix, pages 320-1).

---

*Wold, Herman, in association with Lars Jureen. 1953. Demand Analysis. A Study in Econometrics. John Wiley and Sons, New York. pp. 43-5; pp. 209-13.

APPENDIX B

RUNOFF WATER QUALITY ANALYSIS METHODS

WATKINSVILLE*

Analytical Methodology

Physiochemical Characterization of Soil, Sediment and Runoff--

Runoff samples for sediment analysis were acidified in polyethylene buckets with a few drops H_2SO_4 to about pH 3 to 4 to promote flocculation of the suspended sediment. The clear supernatant was removed and discarded and the sediment was air-dried. The dried sediment was removed, weighed, and stored for later use. The sediment concentration in the original runoff sample was computed knowing the sediment weight and the volume (mass) of the runoff sample.

Particle size distribution, surface area, and organic carbon content were determined on sediment samples from selected runoff events. Similar analyses were also conducted on composite soil samples from each of the sampling areas of the watershed as shown in Figures A2, A4, and A5. Particle size distribution was determined by the hydrometer method,[21] except that dispersion was accomplished using ultrasonic vibration.[22] Organic matter was determined by wet oxidation and potentiometric titration.[23,24] Specific surface area was determined by N_2 gas desorption,[25,26] which measures external surfaces only. This method was chosen as an indicator of total adsorptive capacity because of its rapidity and reproducibility and small sample requirement. Non-expanding clay minerals were predominant in the watershed soils. In preliminary comparative studies of methods, total surface area determined by an ethylene glycol monoethyl ether procedure[27] gave values averaging about three times those of the N_2 desorption procedure.

Pesticide Residue Analysis in Soil, Sediment, and Runoff--

During the project planning stage, it was anticipated that large numbers of runoff and soil core samples would be collected for chemical analysis. After planting, runoff samples were analyzed from each event until the parent pesticide decreased in concentration to a level (depending upon the compound) below the detectable range of the measuring instrument. Each runoff sample received was recorded, a laboratory number assigned, and the samples placed under refrigeration at 4°C pending analysis.

---

*This is a direct excerpt from the following report: Smith, C.N., R.A. Leonard, G.W. Langdale, and G.W. Bailey, 1978. Transport of Agricultural Chemicals from Small Upland Piedmont Watersheds., U.S. EPA (Athens, GA) and U.S.D.A. (Watkinsville, GA) EPA-600/3-78-056 pp. 57-60.

Soil core samples for persistence and mass balance computations were obtained after runoff events. Each core sample was recorded, a laboratory number assigned, and placed in a freezer at -18°C pending analysis. All core data were reported on a moisture free basis.

An analytical method was needed to analyze the parent pesticides in runoff (water and sediment) and soils at a minimum sensitivity in the low parts per million (ppm) for paraquat and the low parts per billion (ppb) for trifluralin and diphenamid. "Production line" analysis was necessary to provide a large sample throughput in a minimum amount of time. In addition, a rapid analytical procedure would reduce the risk of trifluralin loss by volatilization and degradation.

An integrated method fulfilling these requirements was developed.[28,29] This method was later used for the herbicides atrazine, propazine, and cyanazine. These compounds, however, required adjustment of the soil moisture to at least 20 percent to ensure efficient extraction.

2,4-D was analyzed by a modification of the method of Woodham *et al.*[30] as follows: Residues of 2,4-D were determined in soil, sediment, and water by solvent extraction, acidification, and esterification to the methyl ester using diazomethane. The amount of the acid herbicide present was determined by electron capture gas chromatography. A series of 2,4-D-fortified soil and water samples as the free acid were analyzed using the final method. Recoveries run in replicate ranged from 96.7 to 98.4% in soil and water.

Fortified soil and water samples using the 2,4-D formulation (dacamine) consistently ranged from 87 to 91% recovery on duplicates ranging from 2 ppb to 400 ppm. This broad range of levels was run to assure that the length of reaction time of the herbicide with the esterifying reagent and the amount of reagent used would not affect the increase or decrease of the ester recovery. Interferences from soil extractions were eliminated by a H_2O/CH_2Cl_2 shakeout of the acetone/soil extract at the time of acidification. The CH_2Cl_2 extract was evaporated to 1 to 2 ml and transferred to 15-ml conical centrifuge tubes. The remaining CH_2Cl_2 was evaporated just to dryness, and 2 to 3 ml of ether was added at the time of esterification. Fortified samples and 2,4-D standards were run as controls with each set of 20 samples extracted and esterified.

The analyses were performed by using a Tracor MT-220 gas chromatograph equipped with a Coulson electrolytic conductivity detector operating in the nitrogen mode. Colorimetric determinations of paraquat were made by using a Perkin-Elmer Model 202 recording spectrophotometer equipped with an auxiliary recorder and scale expansion accessory.

Chloride and Plant Nutrient in Soil and Runoff--

Soil and runoff samples were stored at -10°C until ready for extraction. Subsamples of unfiltered runoff were stored at -10°C and the remaining subsamples were filtered through a 0.60-μm Nucleopore membrane and the filtrate stored at -10°C. Sediment was not analyzed separately because sediment concentrations in runoff were occasionally so low that collection of an adequate sample by filtration was impractical.

Nutrients were extracted from the frozen soil and runoff samples by placing a 5-gram sample of the frozen material into a 125-ml Erlenmeyer flask with 50 ml of distilled water and shaking the suspension for 1 hour on a wrist-action shaker. Sample weights were corrected for water content from values determined on separate samples taken during the initial sampling (see Table C22 through C35). The suspension was filtered through Whatman Number 41 filter paper, and the filtrate was returned to storage at -10°C until analysis.

Chemical analysis was later accomplished by allowing the frozen test solution to equilibrate to room temperature before proceeding with the selected automated procedures. Technicon auto-analyzer procedures were used exclusively, varying analytical manifold configurations, reaction solutions, and absorption cell lengths as required to give the required sensitivity in the particular colorimetric method.[31-35]

Nitrate-N and Chloride--Nitrates and chloride were determined on a dual channel system using the ferric-mercuric thiocyanate color complex for chlorides and the cadmium reduction procedure for nitrates.

Ortho-phosphorus--Filtered runoff samples were analyzed for ortho-P using the phosphomolybdenum-ascorbic acid blue color complex. Values reported as ortho-P are often referred to as molybdate reaction phosphorus (MRP).

NH_3-N--Ammonia was determined in the filtered and unfiltered runoff samples using the Berthelot color reaction. Differences between the filtered and unfiltered samples are assumed to represent exchangeable NH_3-N and reactive amines displaced from the particulate phase in the alkaline medium.

Total Kjeldahl Nitrogen (TKN)--Filtered and unfiltered runoff samples were predigested in a Technicon BD-40 block digestor with subsequent measurement of the ammonia produced. The quantitation of ammonia was achieved by the Berthelot reaction.

Total Phosphorus--Phosphorus in the filtered runoff samples was hydrolized with ammonium persulfate and sulfuric acid in a pressure cooker at one bar for 30 minutes prior to colorimetric determination of P. The unfiltered samples were digested in a mixture of 1:4 ($HClO_4$:HNO_3) acid until fumes of $HClO_4$ appeared. The residue was then taken up in distilled water and analyzed for total P.

Acid Extractable Phosphorus (Available P)--Available soil P was extracted with a double acid (0.05N HCl in 0.025N H_2SO_4) solution and determined colorimetrically.[36]

Watkinsville Data Base References

21. Day, P.R. Particle Fractionation and Particle Size Analysis. In: Method of Soil Analysis, Part I, Black, C.A. (ed.). Agron. Monograph No. 9.

22. Bouget, S.J. Ultrasonic Vibration for Particle-Size Analysis. Can. J. Soil Sci. 48:372-373. 1968.

23. Jackson, M.L. Soil Chemical Analysis. Englewood Cliffs, Prentiss-Hall, Inc. 1958.

24. Ravek, A. and Y. Aurimelach. Potentiometer Determination of Soil Organic Matter. Soil Sci. Soc. of Amer. Proc. 36:967. 1972.

25. Cremer, E. and H. Huck. Determination of Very Low Surface Area. Institute of Physical Chemistry, University of Innsbruck, Innsbruck, Austria. 1965.

26. Kremen, J., J.S. Lararias, and U.R. Dirtz. Surface Area Determination by Equilibrium Gas Adsorption in Nitrogen-Helium Systems. Rev. of Scientific Instruments. 37:1265-1266. 1966.

27. Neilman, M.D., D.L. Carter, and C.L. Gonzalez. The Ethylene Glycol Monoethyl (EGME) Technique for Determining Soil Surface Area. Soil Sci. 100:409-413. 1965.

28. Payne, W.R., Jr., J.D. Pope, Jr., and J.E. Benner. An Integrated Method for Paraquat, Diphenamid, and Trifluralin in Soil and Runoff from Agricultural Land. J. Agr. Food Chem. 22(1):79-82. 1974.

29. Pope, J.D., Jr., and J.E. Benner. Colorimetric Determination of Paraquat Residues in Soil and Water. J. Assoc. Offic. Anal. Chem. 57(1):202-240. 1974.

30. Woodham, D.W., W.G. Mitchell, C.D. Loftis, and C.W. Collier. An Improved Gas Chromatographic Method for the Analysis of 2,4-D Free Acid in Soil. J. Agr. Food Chem. 19(1):186-188. 1971.

31. Technicon Autoanalyzer Methodology. Individual Simultaneous Determination of Nitrogen and/or Phosphorus in BD Acid Digest. Industrial Method No. 329-74W, Tarrytown, NY. 1975.

32. Technicon Industrial Systems. Chlorides in Water and Wastewater. Industrial Method No. 99-70W, Tarrytown, NY. 1971.

33. Technicon Industrial Systems. Ammonia in Water and Seawater. Industrial Method No. 154-71W, Tarrytown, NY. 1973.

34. Technicon Industrial Systems. Orthophosphate in Water and Seawater. Industrial Method No. 155-71W, Tarrytown, NY. 1973.

35. Jackson, W.A., C.E. Frost, and D.M. Hildreth. Versatile Multirange Analytical Manifold for Automatic Analysis of Nitrate-Nitrogen. Soil Sci. Soc. Amer. Proc. 39:592-593. 1975.

36. Procedures Used by State Soil Testing Laboratories in the Southern Region of the United States. North Carolina State University, Raleigh, NC. Southern Cooperative Series Bulletin No. 190. 23 p.

BUFFALO BILL WATERSHED*

Experimental Method

Since it has been our hypothesis that farm chemicals are attached to silt particles, our sampling program was predicated to reach the stream in the early stages of increased flow. We set up telephone contacts with farmers living on the Buffalo Bill Watershed asking them to advise us when they anticipated sufficient rain to cause significant rise in stream levels. Upon receiving their call, our sampling teams were immediately activated and were sampling the various chosen sampling points in approximately one hour. We set up nine (9) sampling points over the watershed as indicated on the attached map and attempted to sample these nine points three times during a given run-off incident. Approximately one hour was consumed in sampling the nine consecutive sampling points and several of the rainfall run-off incidents started and finished within the space of 4-6 hours. The samples were collected for various parameters including pesticides, nitrogens, solids, fecal coliforms, phosphorus, dissolved oxygen, BOD and pH. These samples were taken back to the laboratory in Iowa City and analyzed by chemists and microbiologists on our regular, full-time staff.

Rainfall data was collected on three (3) sections of the 3500 acre watershed and flow measurements were made on the final downstream aggregate sampling station in Jones Creek (Station #9).

Information regarding the application of farm chemicals was collected during the fall operation and this information appears on the map. Our sampling regimens were designed timewise to show water quality at low-flow-non-runoff periods as well as the degraded water quality during runoff and high turbidity stages.

The various sample stations (1 through 9) were selected to fractionate out the various types of land as well as its usage. Sample sites 1, 2, 3 and 4 were sited to establish runoff water quality characteristics in upland high slope corn and bean fields, while sample site #5 drains essentially pasture land. Sample sites 6, 7 and 8 were chosen because they drain the flatter, low slope corn and bean lands on the downstream part of the watershed. Sample site #9 is the composite of all this drainage and served as the single flow measurement point. We had originally hoped to have flow measurements on more points than just number 9 but the suddenness with which the proposal was funded made the procurement of flow sampling devices impossible.

---

*This is a direct excerpt from the following report: Morris, R.L. and L. Johnson, 1974. Buffalo Bill Watershed Agricultural Runoff Study, Iowa State Hygienic Laboratory, University of Iowa, Iowa City, Iowa. pp. 4-9. Report and data used with kind permission of Dr. Roger Splinter of the Iowa State Hygienic Laboratory. Note that materials presented here provide only limited information about methods but were included for the insights they provide into overall study methods. Figures and tables were not included here.

Rainfall and Flow

Rainfall and flow data are shown in Figure 1 and Table 1 with the turbidity and flow being graphed for station #9. The rainfall data are the average of rainfall collected at the three (3) stations over the watershed. Rainfall data were collected in standard, commercial rainfall gauges and the flow measurements were made according to cross sectional areas and velocity calculations recommended to us by personnel of the U.S. Geological Survey. Because these are not instrumental finite measurements, we have labeled the graph and table as estimated flows. The flow measurements do correlate well with the logic of actually measured rainfall and turbidity so that they appear to have considerable validity.

Perusal of Figure 1 indicates that significant rainfall peaks correlate quite well with increases in estimated flow and that the turbidity values also follow this same flow increase pattern. Absolute conformity of flow and turbidity with rainfall should not be expected because the manner in which the watershed received the rainfall is quite variable. Sometimes an inch of rain falls in a matter of a few minutes and at other periods the one inch rain income occurred over several hours. This variability of rainfall produces different lag periods and peak heights with respect to both flow and turbidity. Figure 1 indicates that there were four (4) significant storm episodes during the five month period. These episodes occurred on August 13th, September 26th, October 10th and December 4th. It will be noted by scrutiny of following data that various levels of rainfall under different conditions produce different levels of silt movement and attendant farm chemicals.

Dissolved and Suspended Solids

Figure 2 and Table 2 contain information correlating turbidity levels and both dissolved and suspended solids. A direct correlation between the suspended or non-filterable solids is shown to a high degree. The filterable or dissolved solids show the same high degree of correlation except in an inverse manner as would be expected. This figure certainly indicates that in a watershed where poor conservation procedures exist, we demonstrate the expected high siltation rates which are so noticeable in midwestern streams. Since the August 13th rain was approximately three inches and the September 26th and October 10th rains were about 0.9 inches, it shows that there is not a direct relationship between the amount of rainfall and the movement of dissolved and suspended solids. Again, this is due in part at least to the duration of rainfall income. Other factors such as plant cover and soil moisture are undoubtedly highly influential in the siltation effect produced by a given amount of rain.

Pesticides

Figure 3 shows the relationship between turbidity and the movement of dieldrin during runoff periods. It should be noted here that aldrin was applied above station 8 or 9 yet the upstream sampling points indicated considerable involvement of dieldrin attached to the moving silt particles. This obviously can be attributed in part to the durability of aldrin applied on the watershed the previous year. Suffice it to say, rises in turbidity produced by the rainfall incidents in every case resulted in significant increases of the dieldrin content. This is in direct agreement with other studies performed by our laboratory in different areas of the state over the past several years. There were no commercial applicators or manufacturing operations associated with this watershed area. The possibility of dieldrin movement by volatility or wind transport from areas outside the Buffalo Bill Watershed was not explored during this five month study, but we certainly intend to review this possibility in the next seven month extended study period. Air samples and rainfall will be analyzed for aldrin-dieldrin content before and after application seasons. Non-durable pesticides were not detected due to the fact that our first samples were taken in mid-August, three months or more subsequent to actual application.

Fecal Coliform

Figure 3 shows the high correlation between turbidity, flow and fecal coliform density. It is interesting to note that the highest levels of fecal coliform at station #9 were not achieved during the greatest rainfall period on August 13th. We have always felt that the earlier part of rainfall produced runoffs probably contained the largest amount of surface leached fecal material and the highest levels were achieved during the two 0.9 inch rains on September 4th and October 10th. Also, it can be noted that the earlier samples taken during the August 13th runoff contained the highest fecal coliform levels which degraded as the runoff incident proceeded.

Nitrogen Series

Figure 4 and Table 2 indicate a distinct, direct correlation between total organic nitrogen and turbidity as well as a consistent but lower level effect for ammonia nitrogen. Nitrate content does not appear to vary in this instance. The ammonia nitrogen levels recorded are within the Lower Water Quality Standard limits and are not nearly as high as we have frequently recorded on other streams within our state. Without having adequate information on the application of commercial and natural fertilizers, we are unable to really determine the source of these nitrogenous materials. It is hoped that the time afforded for this type of application data in the next surveillance period will permit us to render valid judgments on nitrogenous material source. Many of the farmers involved were not exactly sure what they had used on their fields and in what amounts.

Phosphates

Figure 5 shows the relationship of phosphate with respect to turbidity variation and this parameter does not exhibit the correlation found with previously discussed constituents. It does indicate however that phosphorus materials do move off a watershed such as the Buffalo Bill but the variability in phosphate levels is less. As one would obviously expect, pH exhibits an inverse function with respect to runoff due primarily to dilution effects. Elevated pHs are undoubtedly a result of groundwater recharge of the streams and the moving siltation is not sufficiently soluble to offset the effect of volume dilution.

SEATTLE, WASHINGTON

All water quality analyses represent standard procedures indentifiable through EPA STORET codes. The codes for Seattle data are as follows:

STORET code	PARAMETER
665	Total P
10501	OPO_4-P
605	Organic N
610	NH_3-N
630	$(NO_2 + NO_3)$ - N
530	Suspended solids (SS)
70	Turbidity
94	Conductivity
1051	Pb
1052	Zn

HONEY CREEK*

The analytical method used on each parameter sampled at the Sandusky River material transport stations are given below.

Total Phosphorus

Total phosphorus analyses was performed using the automated colorimetric ascorbic acid reduction method for Technicon Autoanalyzer II systems as described in the Methods for Chemical Analysis of Water and Wastes, U.S. Environmental Protection Agency, 1974, beginning on page 256. The persulfate digestion was performed by heating in an autoclave for 30 minutes at 121°C. For samples with high suspended solids, the samples were filtered through a prewashed glass fiber filter (millipore AP2504700) following removal from the autoclave and before cooling. Standard solutions and blanks were included with each batch of samples and underwent the same digestion procedures as the samples. Reported as P.

---

*This is a direct excerpt from: U.S. Army Corps of Engineers (Buffalo, NY), undated. Water Quality Data for Sandusky River Material Transport Stations, Lake Erie Wastewater Management Study. Data collected and analyzed by Heidelburg College, Tiffin, OH. pp. VIII-X.

Dissolved Orthophosphate

The analytical procedures for dissolved orthophosphate were the same as those used for total phosphorus as described above except that sample pretreatment consisted of filtration of the raw sample through a prewashed Millipore HAWP filter. The filtrate was then directly analyzed by the colorimetric procedure cited above. Reported as P.

Residue, Total Non-Refilterable (Suspended Solids)

Suspended solids were analyzed according to the procedures outlined in the Methods for Chemical Analysis of Water and Wastes (U.S.E.P.A., 1974, page 268 and following). A well mixed sample was filtered through a preweighed glass fiber filter, (Reeve Angel 934AH) and the residue retained on the filter was dried to constant weight at 103-105°C. Weighings were done on a Mettler H20T balance with digital readout to the nearest 0.01 milligrams.

Nitrogen, Nitrate-Nitrite

The automated cadmium reduction method was employed as described in the Methods for Chemical Analysis of Water and Wastes (U.S.E.P.A., 1974 page 207 and following). The analysis was run on the same filtrate as used for dissolved orthophosphate. The value reported included both nitrate and nitrite nitrogen.

In this method a filtered sample is passed through a column containing granulated copper-cadmium to reduce nitrate to nitrite. The nitrite (that originally presented plus reduced nitrate) is determined by diazotizing with sulfanilamide and coupling with N-(1-napthyl)-ethylenediamine colorimetrically. Reported as N.

Nitrogen, Ammonia

The automated colorimetric phenate method was employed as described in the above manual (E.P.A., 1974, page 168 and following). The analysis was run on the same filtrate as used for dissolved orthophosphate and nitrate-nitrite. In this method Alkaline phenol and hypochlorite react with ammonia to form indophenol blue that is proportional to the ammonia concentration. The blue color formed is intensified with sodium nitroprusside. Reported as N.

Specific Conductance

Specific conductance was measured using a Barnstead Model PPM 70CB Conductivity Meter and a YSI conductivity cell. Samples were brought to 25°C prior to measurements. Details of the procedure are outlined in Standard Methods for the Examination of Water and Wastewater, 13th Edition, (1971), page 323.

Silica, Dissolved

Silica analysis was performed using the automated method (Technicon Industrial Method #182-72W) on the same filtrates as used for dissolved orthosphate. Reported as SiO_2.

Chloride

Chloride analysis was performed using the procedures outlined in the Methods for Chemical Analysis of Water and Wastes (U.S.E.P.A., 1974, page 31 and following). Thiocyanate ion (SCN) is liberated from mercuric thiocynate, through sequestration of mercury by chloride ion to form unionized mercuric chloride. In the presence of ferric ion, the liberated SCN forms highly colored ferric thiocynate, in concentration proportional to the original chloride concentration. Reported as Cl.

Iron

The analysis of iron was done according to the procedure given in Standard Methods for the Examination of Water and Wastewater, 13th Edition, (1971), with some modification. A Technicon Autoanalysis was utilized according to the automated method (Technicon Industrial Method #109-71W).

Total Kjeldahl Nitrogen

An ultramicro technique was used in the analysis of Total Kjeldahl Nitrogen (TKN). This procedure is outlined in the Ultramicro Semi-Automated Method for the Simultaneous Determination of Total Phosphorus and Total Kjeldahl Nitrogen in Wastewaters (Environmental Science and Technology, October, 1976). The ammonia is analyzed in the range of .05 to 10.0 mg N/l using the indophenol blue method with automated spectrophotometry at the rate of 30 samples per hour. Reported as N.